



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number 147987

To: Sarvamangala Devi
Location: REM 3C18
Art Unit: 1645
Wednesday, March 16, 2005

Case Serial Number: 10/041775

From: Beverly Shears
Location: Remsen Bldg.
RM 1A54
Phone: 571-272-2528

beverly.shears@uspto.gov

Search Notes

Protein Sequence Searches – February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension .rup) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (uniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.



This Page Blank (uspto)

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	950	93.5	689	2	Q53599	Q53599 staphylococ
2	950	93.5	689	2	Q9ZAP5	Q9ZAP5 staphylococ
3	950	93.5	689	2	Q6GFB8	Q6GFB8 staphylococ
4	907	89.3	584	2	Q9XAS5	Q9XAS5 staphylococ
5	888.5	87.5	687	2	Q9S224	Q9S224 staphylococ
6	886.5	87.3	657	2	Q07318	Q07318 staphylococ
7	884.5	87.1	476	2	Q99Q81	Q99Q81 staphylococ
8	881.5	86.8	581	2	Q8NVR4	Q8NVR4 staphylococ
9	814.5	80.2	584	2	Q9K4S8	Q9K4S8 staphylococ
10	801.5	78.9	535	2	Q9ZAJ2	Q9ZAJ2 staphylococ
11	318	31.3	121	1	OMP7 STAAU	P21223 staphylococ
12	210	20.7	144	2	Q8NXX3	Q8NXX3 staphylococ
13	210	20.7	144	2	Q99VA9	Q99VA9 staphylococ
14	210	20.7	144	2	Q7A6G0	Q7A6G0 staphylococ
15	210	20.7	144	2	Q6GNAU5	Q6GNAU5 staphylococ
16	210	20.7	144	2	Q6GTA6	Q6GTA6 staphylococ
17	146.5	14.4	141	2	Q8G961	Q8G961 staphylococ
18	146.5	14.4	141	2	Q99S64	Q99S64 staphylococ
19	146.5	14.4	141	2	Q7A090	Q7A090 staphylococ
20	146.5	14.4	141	2	Q7A483	Q7A483 staphylococ
21	146.5	14.4	141	2	Q6G7B0	Q6G7B0 staphylococ
22	144.5	14.2	141	2	Q6GEM4	Q6GEM4 staphylococ
23	136	13.4	106	2	Q99Q82	Q99Q82 staphylococ
24	135	13.3	96	2	Q8NVR5	Q8NVR5 staphylococ
25	118	11.6	1178	2	Q7PB80	Q7PB80 fuscobacteri
26	118	11.6	1620	2	Q8IIV3	Q8IIV3 plasmodium
27	116	11.4	1792	2	Q8ID94	Q8ID94 plasmodium
28	115	11.3	374	2	Q8DVE1	Q8DVE1 streptococc
29	114.5	11.3	640	1	GYRB SPICI	P34031 spiroplasma
30	114.5	11.3	1272	2	Q9PR01	Q9PR01 ureaplasma
31	113	11.1	1016	2	Q7PDP5	Q7PDP5 plasmodium

```

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1] SEQUENCE FROM N.A.
RP STRAIN=Newman;
RC Hussain S.M., Herrmann M., Heilmann C., Peters G.;
RA Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL; AJ132841; CAA10802.1; -.
DR InterPro; IPR005298; MAP.
DR Pfam; PF03642; MAP; 6.
KW Signal.
FT SIGNAL 1 30 Potential.
FT CHAIN 31 689 MapN protein.
SQ SEQUENCE 689 AA; 76860 MW; CF88802B4F7C9CD CRC64;

Query Match 93.5%; Score 950; DB 2; Length 689;
Best Local Similarity 97.0%; Pred. No. 1e-55;
Matches 191; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 4 SHHHHHSQIPYITVNGTSQNLSSLTENKNOQISYKDIENKVSPLYFNRGISDIDL 63
Db 41 SLHHGYSKIQIPYITVNGTSQNLSSLTENKNOQISYKDIENKVSPLYFNRGISDIDL 100

Qy 64 RLSQAKYTVHFNGTKRVDLKGAIHTADLINTSDIKAISVNVDTKKQVKDKEAKANVQ 123
Db 101 RLSQAKYTVHFNGTKRVDLKGAIHTADLINTSDIKAISVNVDTKKQVKDKEAKANVQ 160

Qy 124 VPYITVNGTSQNLSSLTENKNOQISYKDIENKVSPLYFNRGISDIDL 183
Db 161 VPYITVNGTSQNLSSLTENKNOQISYKDIENKVSPLYFNRGISDIDL 220

Qy 184 NFKNGTKKVIDLKAGIY 200
Db 221 NFKNGTKKVIDLKAGIY 237

RESULT 3
O6GFB8 PRELIMINARY; PRT; 689 AA.
AC O6GFB8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE MHC class II analog.
GN OrderedLocusNames=SA2030;
OS Staphylococcus aureus (strain MRS4252).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282458;
RN [1] SEQUENCE FROM N.A.
RP PubMed=15213324; DOI=10.1073/pnas.0402521101;
RX Holden M.T.G., Fell E.J., Lindsey J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
RA Felwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,
RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K.M., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL; EX571856; CAG41016.1; -.
DR InterPro; IPR005298; MAP.
DR Pfam; PF03642; MAP; 6.
KW Complete proteome.
SQ SEQUENCE 689 AA; 76848 MW; 4141D8900549099D CRC64;

Query Match 93.5%; Score 950; DB 2; Length 689;
Best Local Similarity 97.0%; Pred. No. 1e-55;
Matches 191; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 4 SHHHHHSQIPYITVNGTSQNLSSLTENKNOQISYKDIENKVSPLYFNRGISDIDL 63
Db 41 SLHHGYSKIQIPYITVNGTSQNLSSLTENKNOQISYKDIENKVSPLYFNRGISDIDL 100

Qy 64 RLSQAKYTVHFNGTKRVDLKGAIHTADLINTSDIKAISVNVDTKKQVKDKEAKANVQ 123
Db 101 RLSQAKYTVHFNGTKRVDLKGAIHTADLINTSDIKAISVNVDTKKQVKDKEAKANVQ 160

Qy 124 VPYITVNGTSQNLSSLTENKNOQISYKDIENKVSPLYFNRGISDIDL 183
Db 161 VPYITVNGTSQNLSSLTENKNOQISYKDIENKVSPLYFNRGISDIDL 220

Qy 184 NFKNGTKKVIDLKAGIY 200
Db 221 NFKNGTKKVIDLKAGIY 237

RESULT 4
O9XAS5 PRELIMINARY; PRT; 584 AA.
AC O9XAS5;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Map-7 protein precursor.
GN Name=map-7;
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1] SEQUENCE FROM N.A.
RA Hussain Shaikh M., Heilmann C., Peters G., Herrmann M.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ243790; CAB50920.1; -.
DR InterPro; IPR005298; MAP.
DR Pfam; PF03642; MAP; 5.
KW Signal.
FT SIGNAL 1 30 Potential.
FT CHAIN 31 584 Potential.
SQ SEQUENCE 584 AA; 65173 MW; 0A39383C55F639C1 CRC64;

Query Match 89.3%; Score 907; DB 2; Length 584;
Best Local Similarity 93.4%; Pred. No. 6.5e-53;
Matches 184; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 4 SHHHHHSQIPYITVNGTSQNLSSLTENKNOQISYKDIENKVSPLYFNRGISDIDL 63
Db 41 SLHHGYSKIQIPYITVNGTSQNLSSLTENKNOQISYKDIENKVSPLYFNRGISDIDL 100

Qy 64 RLSQAKYTVHFNGTKRVDLKGAIHTADLINTSDIKAISVNVDTKKQVKDKEAKANVQ 123
Db 101 RLSQAKYTVHFNGTKRVDLKGAIHTADLINTSDIKAISVNVDTKKQVKDKEAKANVQ 160

Qy 124 VPYITVNGTSQNLSSLTENKNOQISYKDIENKVSPLYFNRGISDIDL 183
Db 161 VPYITVNGTSQNLSSLTENKNOQISYKDIENKVSPLYFNRGISDIDL 220

Qy 184 NFKNGTKKVIDLKAGIY 200
Db 221 NFKNGTKKVIDLKAGIY 237

RESULT 5
Q9S2Z4 PRELIMINARY; PRT; 687 AA.
AC Q9S2Z4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cell surface protein map-w precursor.
GN Name=map-w;
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

```



```

OX NCBI_TaxID=1280;
RN [1]_TaxID=1280;
RP SEQUENCE FROM N.A.
RA STRAIN=wood 46;
RC Hussein Shaikh M., Heilmann C., Peters G., Herrmann M.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ245439; CAB51807.1; -.
DR InterPro; IPR005298; MAP.
DR Pfam; PF03642; MAP; 6.
KW SIGNAL.
FT SIGNAL.
SQ SEQUENCE 687 AA; 77072 MW; 91D429532DB7DD2B CRC64;

Query Match 87.5%; Score 888.5; DB 2; Length 687;
Best Local Similarity 91.3%; Pred. No. 1.4e-51;
Matches 178; Conservative 7; Mismatches 9; Indels 1; Gaps 1;

QY 6 HHHHGSQIPYITVNGTSQNTLSSTFNKNQOISYKDIEKNVSVLYFNRGISDIDLRL 65
DB 43 HHGHSNIQIPYITVNGTSQNTLSSTFNKNQOISYKDIEKNVSVLYFNRGISDIDLRL 102
QY 66 SKQAKYTVHFKNKGRVVDLKGAGIHTADLINTSDIKAISVNVDTKKQVKDKAKANVQVP 125
DB 103 SKQAEYTVHFKNKGRVVDLKGAGIHTADLINTSDIKAISVNVDTKKQVKDKAKANVQVP 161
QY 126 YTIITVNGTSQNTLSNLTFFKNQOISYKOLENNVSVLKNRGITDVLRLSKQAKFTVNF 185
DB 162 YTIITVNGTSQNTLSNLTFFKNQOISYKOLENNVSVLKNRGITDVLRLSKQAKFTVNF 221
QY 186 KNGTKKVIDLKAGIY 200
DB 222 KNGTKKVIDLKAGIY 236

RESULT 6
O07318
ID O07318 PRELIMINARY; PRT; 657 AA.
AC O07318;
DT 01-JUL-1997 (TReMBLrel. 04, Created)
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Outer surface binding 70kD protein (Fragment).
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]_TaxID=1280;
RP SEQUENCE FROM N.A.
RA Yousef Y., Draeger R., Schiltz M., Peter H., Schlesier M.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; X10419; CAAY71446.1; -.
DR InterPro; IPR005298; MAP.
DR Pfam; PF03642; MAP; 6.
FT NON_TER.
SQ SEQUENCE 657 AA; 74044 MW; 1A7980DF1B7E09A3 CRC64;

Query Match 87.3%; Score 886.5; DB 2; Length 657;
Best Local Similarity 90.8%; Pred. No. 1.8e-51;
Matches 177; Conservative 8; Mismatches 9; Indels 1; Gaps 1;

QY 6 HHHHGSQIPYITVNGTSQNTLSSTFNKNQOISYKDIEKNVSVLYFNRGISDIDLRL 65
DB 13 HHGHSNIQIPYITVNGTSQNTLSSTFNKNQOISYKDIEKNVSVLYFNRGISDIDLRL 72
QY 66 SKQAKYTVHFKNKGRVVDLKGAGIHTADLINTSDIKAISVNVDTKKQVKDKAKANVQVP 125
DB 73 SKQAEYTVHFKNKGRVVDLKGAGIHTADLINTSDIKAISVNVDTKKQVKDKAKANVQVP 131
QY 126 YTIITVNGTSQNTLSNLTFFKNQOISYKOLENNVSVLKNRGITDVLRLSKQAKFTVNF 185
DB 132 YTIITVNGTSQNTLSNLTFFKNQOISYKOLENNVSVLKNRGITDVLRLSKQAKFTVNF 191
QY 186 KNGTKKVIDLKAGIY 200
```

```

DB 192 KNGTKKVIDLKAGIY 206

RESULT 7
Q99QS1
ID Q99QS1 PRELIMINARY; PRT; 476 AA.
AC Q99QS1;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Truncated map-w protein.
GN OrderedLocusNames=SAV1938;
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=159878;
RN [1]_TaxID=159878;
RP SEQUENCE FROM N.A.
RC STRAIN=Mu50 / ATCC 700699;
RX MEDLINE=21311952; PubMed=11418146; DOI=10.1016/S0140-6736(00)04403-2;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,
"Whole genome sequencing of methicillin-resistant Staphylococcus
aureus";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003364; BAB58100.1; -.
DR InterPro; IPR005298; MAP.
DR Pfam; PF03642; MAP; 4.
KW Complete proteome.
SQ SEQUENCE 476 AA; 53377 MW; 04E6B5C9DBD782E8 CRC64;

Query Match 87.1%; Score 884.5; DB 2; Length 476;
Best Local Similarity 90.8%; Pred. No. 1.7e-51;
Matches 177; Conservative 7; Mismatches 10; Indels 1; Gaps 1;

QY 6 HHHHGSQIPYITVNGTSQNTLSSTFNKNQOISYKDIEKNVSVLYFNRGISDIDLRL 65
DB 43 HHGHSNIQIPYITVNGTSQNTLSSTFNKNQOISYKDIEKNVSVLYFNRGISDIDLRL 102
QY 66 SKQAKYTVHFKNKGRVVDLKGAGIHTADLINTSDIKAISVNVDTKKQVKDKAKANVQVP 125
DB 103 SKQAEYTVHFKNKGRVVDLKGAGIHTADLINTSDIKAISVNVDTKKQVKDKAKANVQVP 161
QY 126 YTIITVNGTSQNTLSNLTFFKNQOISYKOLENNVSVLKNRGITDVLRLSKQAKFTVNF 185
DB 162 YTIITVNGTSQNTLSNLTFFKNQOISYKOLENNVSVLKNRGITDVLRLSKQAKFTVNF 221
QY 186 KNGTKKVIDLKAGIY 200
DB 222 KNGTKKVIDLKAGIY 236

RESULT 8
Q8NVR4
ID Q8NVR4 PRELIMINARY; PRT; 581 AA.
AC Q8NVR4;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Truncated cell surface protein map-w.
GN Name=truncated map-w; OrderedLocusNames=MW1880;
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]_TaxID=196620;
RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
RX MEDLINE=22040717; PubMed=12044378; DOI=10.1016/S0140-6736(02)08713-5;
```

```
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwano N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA";
RL Lancet 359:1819-1827(2002).
DR EMBL; AP004828; BAB95745.1; -.
DR InterPro; IPR005298; MAP.
DR Pfam; PF03642; MAP; 5.
KW Complete proteome.
SQ SEQUENCE 581 AA; 65095 MW; 912F70CDAD8B8C68 CRC64;

Query Match 86.8%; Score 881.5; DB 2; Length 581;
Best Local Similarity 90.3%; Pred. No. 3.3e-51;
Matches 176; Conservative 8; Mismatches 10; Indels 1; Gaps 1;

QY 6 HHHHGGGQIPYITVNGTSQNLSTFNKNQOISYKDIENKVSVLVFNRGISDIDLRL 65
DB 43 HGHSHNIQIPYITVNGTSQNLSTFNKNQOISYKDIENKVSVLVFNRGISDIDLRL 102
QY 66 SKQAKYTVHFKNKGRVVDLKGAGHTADLINTSDIKAISVNVDTKKQVKDKEAKANQVP 125
DB 103 SKQAEYTVHFKNKGRVVDLKGAGHTADLINTSDIKAISVNVDTKKQVKDKEAKANQVP 161
QY 126 YTIITVNGTSQNLSTFNKNQOISYKDIENKVSVLVFNRGITDVLRLSKQAKFTVNF 185
DB 162 YTIITVNGTSQNLSTFNKNQOISYKDIENKVSVLVFNRGITDVLRLSKQAKFTVNF 221

QY 186 KNGTKKVIDLKAGIY 200
DB 222 KNGTKKVIDLKSGIY 236

RESULT 9
Q9K4S8
ID Q9K4S8 PRELIMINARY; PRT; 584 AA.
AC Q9K4S8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Map-ND2C protein precursor.
GN Name=map-ND2C;
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Newman;
RX MEDLINE=21546656; PubMed=11687475;
RA Hussain M., Becker K., von Eiff C., Peters G., Herrmann M.;
RT "Analogs of Sap protein are conserved and prevalent in clinical
RT Staphylococcus aureus isolates.";
RL Clin. Diagn. Immunol. 8:1271-1276(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Newman D2C ATCC 25904;
RA Hussain Shaikh M.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ290973; CAB94853.1; -.
DR InterPro; IPR005298; MAP.
DR Pfam; PF03642; MAP; 5.
KW Signal.
FT SIGNAL 1 30 Potential.
FT CHAIN 31 584 Map-ND2C protein.
SQ SEQUENCE 584 AA; 65542 MW; EFE45AC0C9615ED6 CRC64;

Query Match 80.2%; Score 814.5; DB 2; Length 584;
Best Local Similarity 81.8%; Pred. No. 1e-46;
Matches 162; Conservative 19; Mismatches 16; Indels 1; Gaps 1;

QY 4 SHHHHGGGQIPYITVNGTSQNLSTFNKNQOISYKDIENKVSVLVFNRGISDIDL 63
DB 41 SLHHGYSKVVHPYAITVNGTSQNLSTFNKNQOISYKDIENKVSVLVFNRGISDIDL 100
```

```
QY 64 RLSQAKYTVHFKNKGRVVDLKGAGHTADLINTSDIKAISVNVDTKKQVKDKEA-KANV 122
DB 101 RLSQAKYTVHFKNKGRVVDLKGAGHTADLINTSDIKAISVNVDTKKQVKDKEA-KANV 160
QY 123 QVPYITVNGTSQNLSTFNKNQOISYKDIENKVSVLVFNRGITDVLRLSKQAKFT 182
DB 161 QVPYITVNGTSQNLSTFNKNQOISYKDIENKVSVLVFNRGITDVLRLSKQAKFT 220
QY 183 VNFNKGTKKVIDLKAGIY 200
DB 221 VNFNKGTKKVIDLKSGIY 238

RESULT 10
Q9Z4J2
ID Q9Z4J2 PRELIMINARY; PRT; 535 AA.
AC Q9Z4J2;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Map protein (Fragment).
GN Name=map;
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Newman;
RA Kreikemeyer B., McDevitt D., Kapur V., Hook M.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ223806; CAA11555.1; -.
DR InterPro; IPR005298; MAP.
DR Pfam; PF03642; MAP; 5.
FT NON_TER 1 535 Map protein.
FT CHAIN <1 535
SQ SEQUENCE 535 AA; 60459 MW; D2566G318AC64F2A1 CRC64;

Query Match 78.9%; Score 801.5; DB 2; Length 535;
Best Local Similarity 84.6%; Pred. No. 7.1e-46;
Matches 159; Conservative 19; Mismatches 9; Indels 1; Gaps 1;

QY 14 IPYITVNGTSQNLSTFNKNQOISYKDIENKVSVLVFNRGISDIDLRLSKQAKYTV 73
DB 2 VPYAITVNGTSQNLSTFNKNQOISYKDIENKVSVLVFNRGISDIDLRLSKQAKYTV 61
QY 74 HFKNGTKRVDLKGAGHTADLINTSDIKAISVNVDTKKQVKDKEA-KANVQVPYITVNG 132
DB 62 YFKNKGRVVDLKGAGHTADLINTSDIKAISVNVDTKKQVKDKEA-KANVQVPYITVNG 121
QY 133 TSQNLSTFNKNQOISYKDIENKVSVLVFNRGITDVLRLSKQAKFTVNFNKGTKKV 192
DB 122 TSQNLSTFNKNQOISYKDIENKVSVLVFNRGITDVLRLSKQAKFTVNFNKGTKKV 181
QY 193 IDLKAGIY 200
DB 182 IDLKSGIY 189

RESULT 11
OMP7_STAAU
ID OMP7_STAAU STANDARD; PRT; 121 AA.
AC P21223;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE 70 kDa outer membrane protein precursor (Fragment).
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COL;
```

```

RX MEDLINE=89263748; PubMed=2726469;
RA Projan S.J., Kornblum J., Kreiswirth B., Moghazeh S.L., Eisner W.,
RA Novick R.P.;
RT "Nucleotide sequence: the beta-hemolysin gene of Staphylococcus
RT aureus.";
RL Nucleic Acids Res. 17:3305-3305(1989).
RN [2]
RP SEQUENCE OF 31-52.
RC STRAIN=ATCC 25923;
RA Yousef Y., Schiltz E., Vogt A.;
RL Submitted (MAR-1991) to Swiss-Prot.
RN [3]
RP SEQUENCE OF 31-53.
RC STRAIN=Newman;
RA Hussein M.S., Herrmann M., Chhatwal G.S., Peters G.;
RT "Map-n protein deficient mutant of Staphylococcus aureus Newman binds
RT less to fibronectin and collagen coated surfaces. A 70 kDa vitronectin
RT protein of Staphylococcus aureus Newman.";
RL Submitted (FEB-1999) to Swiss-Prot.
CC -!- FUNCTION: Binds various plasma and ECM-proteins.
CC -!- SUBCELLULAR LOCATION: Outer membrane.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X13404; CAA31768.1; -.
DR PIR; S15765; S15765.
DR InterPro; IPR005298; MAP.
DR Pfam; PF03642; MAP; 1.
KW Direct protein sequencing; Outer membrane; Signal.
FT SIGNAL 1 30
FT CHAIN 31 >121 70 kDa outer membrane protein.
FT VARIANT 111 111 Y -> S (in strain UM4102).
FT VARIANT 111 111 Y -> F (in strain UM4082).
FT NON_TER 131 121
SQ SEQUENCE 121 AA; 13240 MW; DCFE33E806241E4A CRC64;

Query Match 31.3%; Score 318; DB 1; Length 121;
Best Local Similarity 76.5%; Pred. No. 3.7e-14;
Matches 62; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 4 SHHHHHSQIPYITVNGTSQNLISLTFFNKNQOISYKDIENKVKSVLYFNRGISDIDL 63
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 64 RLSKOAKTVVHPKNGTKRVVD 84
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 101 RLSKOAKTVVHPKNGTKKVID 121
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 12
QY Q8NKE3 PRELIMINARY; PRT; 144 AA.
AC Q8NKE3;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE MW0863 protein.
GN OrderedLocusNames=MW0863;
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
RX MEDLINE=22040717; PubMed=12044378; DOI=10.1016/S0140-6736(02)08713-5;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,

```

```

RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL; AP004825; BAB94728.1; -.
DR InterPro; IPR005298; MAP.
DR Pfam; PF03642; MAP; 1.
KW Complete proteome.
SQ SEQUENCE 144 AA; 15838 MW; F29A2AE30ECD4563 CRC64;

Query Match 20.7%; Score 210; DB 2; Length 144;
Best Local Similarity 44.3%; Pred. No. 8.1e-07;
Matches 43; Conservative 22; Mismatches 32; Indels 0; Gaps 0;

QY 14 IPYITVNGTSQNLISLTFFNKNQOISYKDIENKVKSVLYFNRGISDIDLRLSKOAKTV 73
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 74 HFKNGTKRVDLKGAIHTADLINTSDIKAISVNVDTK 110
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 108 TWKDGSKKEVDLKKDSYTANLFDNSIKQIDINVTK 144
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 13
QY Q99VA9 PRELIMINARY; PRT; 144 AA.
AC Q99VA9;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=SAV0981;
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MU50 / ATCC 700699;
RX MEDLINE=21311952; PubMed=11418146; DOI=10.1016/S0140-6736(00)04403-2;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003360; BAB57143.1; -.
DR PIR; F89865; F89865.
DR InterPro; IPR005298; MAP.
DR Pfam; PF03642; MAP; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 144 AA; 15898 MW; ABFE3FE30ED9506C CRC64;

Query Match 20.7%; Score 210; DB 2; Length 144;
Best Local Similarity 44.3%; Pred. No. 8.1e-07;
Matches 43; Conservative 22; Mismatches 32; Indels 0; Gaps 0;

QY 14 IPYITVNGTSQNLISLTFFNKNQOISYKDIENKVKSVLYFNRGISDIDLRLSKOAKTV 73
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 74 HFKNGTKRVDLKGAIHTADLINTSDIKAISVNVDTK 110
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 108 TWKDGSKKEVDLKKDSYTANLFDNSIKQIDINVTK 144
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 14
QY Q7A6G0 PRELIMINARY; PRT; 144 AA.
AC Q7A6G0;

```



```

Qy 1 MRGSHHHHGGSQIPYTTITVNGTSQNILSLTFNKNQOISYKDIENTKVKSVLYFNRGISD 60
Db 41 MDGKH-----TVPTIISVDGITALHRTYFVFPENKVKVLYQIEDSKVKNELASQRGVTT 93

Qy 61 IDRLSKQAKYTVHFKNKTRKRVVDLKAGIHTADLINTSDIKAISVNV 107
Db 94 EKINNAQTATYTLTLDNGKNKVVNLKKNDDAKNSIDPSTIKQIQIVV 140

RESULT 21
Q6G7B0 PRELIMINARY; PRT; 141 AA.
AC Q6G7B0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Putative exported protein.
GN OrderedLocusNames=SAS2104;
OS Staphylococcus aureus (strain MSSA476).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282459;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15213324; DOI=10.1073/pnas.0402521101;
RA Holden M.T.G., Feil E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
RA Feltwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,
RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,
RA Mouton R.P., Parkhill J.,
RA Peacock S., Simmonds M., Stevens K., Whitehead S., Barrall B.G.,
RA Sharp S., Smeaton M.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL; BX571857; CAG43913.1; -.
DR InterPro; IPR005298; MAP.
DR Pfam; PF03642; MAP; 1.
KW Complete proteome.
SQ SEQUENCE 141 AA; 15426 MW; 42142DB720E8DE51 CRC64;

Query Match 14.4%; Score 146.5; DB 2; Length 141;
Best Local Similarity 31.8%; Pred. No. 0.014;
Matches 34; Conservative 26; Mismatches 40; Indels 7; Gaps 1;

Qy 1 MRGSHHHHGGSQIPYTTITVNGTSQNILSLTFNKNQOISYKDIENTKVKSVLYFNRGISD 60
Db 41 MDGKH-----TVPTIISVDGITALHRTYFVFPENKVKVLYQIEDSKVKNELASQRGVTT 93

Qy 61 IDRLSKQAKYTVHFKNKTRKRVVDLKAGIHTADLINTSDIKAISVNV 107
Db 94 EKINNAQTATYTLTLDNGKNKVVNLKKNDDAKNSIDPSTIKQIQIVV 140

RESULT 22
Q6GEM4 PRELIMINARY; PRT; 141 AA.
AC Q6GEM4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Putative exported protein.
GN OrderedLocusNames=SAR2295;
OS Staphylococcus aureus (strain MRSA252).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282458;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15213324; DOI=10.1073/pnas.0402521101;
RA Holden M.T.G., Feil E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,

```

```

RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
RA Feltwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,
RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,
RA Mouton R.P., Parkhill J.,
RA Peacock S., Simmonds M., Stevens K., Whitehead S., Barrall B.G.,
RA Sharp S., Quail M.A., Rabinowitsch E., Rutherford K.M., Sanders M.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL; BX571856; CAG41274.1; -.
DR InterPro; IPR005298; MAP.
DR Pfam; PF03642; MAP; 1.
KW Complete proteome.
SQ SEQUENCE 141 AA; 15479 MW; EC4882DA8377A605 CRC64;

Query Match 14.2%; Score 144.5; DB 2; Length 141;
Best Local Similarity 31.8%; Pred. No. 0.02;
Matches 34; Conservative 26; Mismatches 40; Indels 7; Gaps 1;

Qy 1 MRGSHHHHGGSQIPYTTITVNGTSQNILSLTFNKNQOISYKDIENTKVKSVLYFNRGISD 60
Db 41 MDGKH-----TVPTIISVDGITALHRTYFVFPENKVKVLYQIEDSKVKNELASQRGVTT 93

Qy 61 IDRLSKQAKYTVHFKNKTRKRVVDLKAGIHTADLINTSDIKAISVNV 107
Db 94 EKINNAQTATYTLTLDNGKNKVVNLKKNDDAKNSIDPSTIKQIQIVV 140

RESULT 23
Q99Q82 PRELIMINARY; PRT; 106 AA.
AC Q99Q82;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Truncated map-w protein.
GN OrderedLocusNames=SAV1937;
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Mu50 / ATCC 700699;
RX MEDLINE=21311952; PubMed=11418146; DOI=10.1016/S0140-6736(00)04403-2;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Murayama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogatawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003364; BAB58099.1; -.
DR PIR; H89982; H89982.
DR InterPro; IPR005298; MAP.
DR Pfam; PF03642; MAP; 1.
KW Complete proteome.
SQ SEQUENCE 106 AA; 12260 MW; 605482EDD4CC2200 CRC64;

Query Match 13.4%; Score 136; DB 2; Length 106;
Best Local Similarity 27.1%; Pred. No. 0.053;
Matches 26; Conservative 27; Mismatches 43; Indels 0; Gaps 0;

Qy 12 SQIPYTTITVNGTSQNILSLTFNKNQOISYKDIENTKVKSVLYFNRGISDILRLSKQAKY 71
Db 9 NVKFPVTKNFENIVSNFVFNASKITINDUSIKLSAWANDQGITKHDIGLAERAVY 68

Qy 72 TVHFKNKTRKRVVDLKAGIHTADLINTSDIKAISVNV 107
Db 69 KYVFNKSSKYVDLKTVEYKDERVFKATDIKKVDIEL 104

```

```

RESULT 24
QBNVRS
ID QBNVRS PRELIMINARY; PRT; 96 AA.
AC QBNVRS5;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Truncated map-w protein.
GS Name=truncated map-w; OrderedLocusNames=MW1879;
OS Staphylococcus aureus (strain MW2);
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=156620;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
RX MEDLINE=22040717; PubMed=12044378; DOI=10.1016/S0140-6736(02)08713-5;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL: AP004828; BAB95744.1; -.
DR InterPro; IPR005298; MAP.
DR Pfam; PF03642; MAP; 1.
KW Complete proteome.
SQ SEQUENCE 96 AA; 11160 MW; DA9121B77E64FA51 CRC64;

Query Match 13.3%; Score 135; DB 2; Length 96;
Best Local Similarity 28.3%; Pred. No. 0.055;
Matches 26; Conservative 25; Mismatches 41; Indels 0; Gaps 0;

QY 16 YTTVNGTSQNILSLTFNKNQOISYKDIENKVKSVLYFNRGISDIDLRLSKOAKYTVHF 75
DB 3 FVINKENIVNEFVYNASKITINDLSKLKSAIANDQOITKHDIELAEAVKYVF 62

QY 76 KNGTKRVVDLKAGIHTADLINTSDIKASVNV 107
DB 63 KNGSSKYVDLKTVEKDERVFRATDICKVDIEL 94

RESULT 25
Q7P8M0 PRELIMINARY; PRT; 1178 AA.
AC Q7P8M0;
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DE Hypothetical protein.
GS Name=FN2276;
OS Fusobacterium nucleatum subsp. vincentii ATCC 49256.
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=209882;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 49256;
RA Kapatral V., Ivanova N., Anderson I., Reznik G., Bhattacharyya A.,
RA Gardner W.L., Mikhailova N., Larsen N., D'Souza M., Walunas T.,
RA Haseelkorn R., Overbeek R., Kyrpides N.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABF01000001; EAA25241.1; -.
KW Hypothetical protein.
SQ SEQUENCE 1178 AA; 134211 MW; 8C143D8C030F952D CRC64;

Query Match 11.6%; Score 118; DB 2; Length 1178;
Best Local Similarity 22.2%; Pred. No. 13;
Matches 47; Conservative 35; Mismatches 70; Indels 60; Gaps 7;

QY 18 ITVNGTSQNILSLTFNKNQOISYKDIENK-VKSVLYF--NRGISDIDLRLSKOAKYTVH 74
DB 2 ISFDRVDENILNKIPIREKKDLKXINIKDKITIVHYSNDRGLS----- 46

QY 75 FNGTKRVVDLKAGIHTADLINTSDIKASVNVVDTKQVKQKAEKANQVPTIYVNGTS 134
DB 47 -----IKTTMKPNDSEFGIELN-DPNLYVSSKGGKRNLSARILTKVKGIP 91

QY 135 QNLSNLTFKKNQ-----QISYKDLNNVKSVLKSN--RGITDVLRL 175
DB 92 ENIALSVENQKNDTIILALKSPIKNDIPIINIRGKIENQKDLKANIDSIVDFNWDY 151

QY 176 SKQA-----KFTVNFNKGTKKVIDLKAGI 199
DB 152 KDKKKLAKIYGNKFTINVDVKKLTDCGRI 183

RESULT 26
Q8IIV3 PRELIMINARY; PRT; 1620 AA.
AC Q8IIV3;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Hypothetical protein.
GS ORFNames=PF11_0060;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255705; PubMed=12368864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Perteira M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Nature 419:498-511(2002).
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AE014836; AAN35649.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR ProDom; PD000001; Prot_kinase; 2.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Hypothetical protein; Kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1620 AA; 195179 MW; CB758269933C016 CRC64;

Query Match 11.6%; Score 118; DB 2; Length 1620;
Best Local Similarity 23.8%; Pred. No. 19;
Matches 46; Conservative 40; Mismatches 81; Indels 26; Gaps 9;

QY 9 HGSQIPYTTIVNGTSQNILSLTFNKNQOISYKDIENKVKSVLYFNRGISDIDLRLSK 67
DB 1158 YHPSNGNIFITTHLSNTYSN---ERFIELFVQKKKKILSFHNSFEFFIINN 1214

QY 68 QAKYTVHFNGTKRVVDLKAGIHTADLINTS-DIKASVNVVDTKQVKD-----KE 117
DB 1215 KIQNQYIQGNLHLQIKKDIHS---INKELDIKKKNENQOQNKQKNQYVYHQ 1271

```


QY	118	AKANVQVPYITVNGTSONILSNLTFKKNQOISYKDLNNVKSVLKSNRGITDVLBSK	177
Db	1272	NOTNLXP-NIMVKGKQN--ENESEYKKNNEIKYKNEENNINTIQHNNHNI-----QIYK	1324
QY	178	QAKFTVNFKNGTK	190
Db	1325	DKR--INFMMPHK	1335
RESULT 27			
Q81D94	Q81D94 PRELIMINARY; PRT; 1792 AA.		
ID	Q81D94	PRELIMINARY; PRT; 1792 AA.	
AC	Q81D94;		
DT	01-MAR-2003	(TrEMBLrel. 23, Created)	
DT	01-MAR-2003	(TrEMBLrel. 23, Last sequence update)	
DT	01-MAR-2003	(TrEMBLrel. 23, Last annotation update)	
DE	Malaria antigen.		
GN	Name=MAL13P1.304;		
OS	Plasmodium falciparum (isolate 3D7).		
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
OX	NCBI_TaxID=36329;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,		
RA	Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,		
RA	Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.;		
RL	Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.		
RL	EMBL; AL844509; CAD52732.1; -.		
DR	InterPro; IPRO11591; Botulinum.		
DR	ProDom; PD001963; Botulinum; 1.		
SQ	SEQUENCE 1792 AA; 210163 MW; 4959B52B59D3794 CRC64;		

[illegible]

```

RESULT 28
Q8DVE1
ID Q8DVE1 PRELIMINARY; PRT; 374 AA.
AC Q8DVE1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative cell division protein FtsQ (DivIB).
GN Name=ftsQ; OrderedLocusNames=SMU.550;
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_taxid=1309;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype c;
RX MEDLINE=22295063; PubMed=1239718; DOI=10.1073/pnas.172501299;
RA Ajdic D.G., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H.G., Lin S.P.,
RA Qian Y., Li S., Zhu H., Najjar F.Z., Lai H., White J., Roe B.A.,

```

[illegible]

RESULT 29

GYRB_SPICI STANDARD; PRT; 640 AA.
 ID GYRB_SPICI AC P34031;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE DNA gyrase subunit B (EC 5.99.1.3).
 GN Name-gyrb;
 OS Spiroplasma citri.
 OC Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;
 OC Spiroplasmataceae; Spiroplasma.
 OX NCBI_TaxID=2133;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=R8A2HP;
 RX MEDLINE=94230261; PubMed=7764984;
 RA Ye F., Renaudin J., Bove J.M., Laigret F.;
 RT "Cloning and sequencing of the replication origin (oriC) of the
 RT Spiroplasma citri chromosome and construction of autonomously
 RT replicating artificial plasmids.";
 RL Curr. Microbiol. 29:23-29(1994).
 CC -1- FUNCTION: DNA gyrase negatively supercoils closed circular double-
 CC stranded DNA in an ATP-dependent manner and also catalyzes the
 CC interconversion of other topological isomers of double-stranded
 CC DNA rings, including catenanes and knotted rings.
 CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
 CC of double-stranded DNA.
 CC -1- SUBUNIT: Made up of two chains. The A chain is responsible for DNA
 CC breakage and rejoining; the B chain catalyzes ATP hydrolysis. The
 CC enzyme forms an A2B2 tetramer.
 CC -1- SIMILARITY: Belongs to the type II topoisomerase family.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; Z19108; CAA79523.1; -.


```

DR PIR: S35734; S35734.
DR HSP: P06982; 1A26.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR002288; DNA_gyraseB_C.
DR InterPro: IPR000565; DNA_gyrB.
DR InterPro: IPR001241; DNA_topoisom.
DR InterPro: IPR006171; Toprim_dom.
DR Pfam: PF00204; DNA_gyraseB_1.
DR Pfam: PF00986; DNA_gyraseB_C_1.
DR Pfam: PF02518; HATPase_C_1.
DR Pfam: PF01751; Toprim_1.
DR PRINTS: PR00418; TP12FAMILY.
DR ProDom: PD149633; DNA_gyrase_B; 1.
DR SMART: SM00387; HATPase_C_1.
DR SMART: SM00433; TOP2C; 1.
DR TIGRFAMs: TIGR01059; GyrB; 1.
DR PROSITE: PS00177; TOPOISOMERASE II; 1.
DR ATP-binding; Isomerase; Topoisomerase.
SQ SEQUENCE 640 AA; 72763 MW; A7403DE2B098D25A CRC64;

Query Match 11.3%; Score 114.5; DB 1; Length 640;
Best Local Similarity 24.5%; Pred. No. 12;
Matches 52; Conservative 50; Mismatches 61; Indels 49; Gaps 14;

QY 9 HHGQIPYITVNGTSQNILSLTFNKQOI-----SYKDIEKNKSVLYFNRG--IS 59
DQ 145 HGGQILPTIKVGSTGTITVFLPDEKIPKETTIFSTIQNRKQLVFLNKGLEIS 204
QY 60 DIDRLSKQAKYTVH-FKNGTK-RVVDLKAGIHTADLINTSDIKAKAISVNVDTKKQVKKE 117
DQ 205 LVDLREDEEKVLFQFNGIKDYVLELNKTICTP-----LNDVFVY-----EGIED-- 251
QY 118 AKANVOQPYTITVNGT-SQNILS-----TFKNQOQIS-YKDLENNVKS-VLKS 165
DQ 252 ---NIWVEFGLOYNDYSENIFSCFNININTEGGTHEGALRAIVREINNVKFNQINKON 308
QY 166 RGITVDLRLSKQAKET-VNFKNGCTKVIDLK 196
DQ 309 KGNED-----KFTWDDIKEGMTIISIR 331

RESULT 30
QSPR01 PRELIMINARY; PRT; 1272 AA.
ID Q9PR01
AC Q9PR01;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Conserved hypothetical.
GN OrderedLocusNames=U0143;
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.
OX NCBI_TaxID=134821;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 700970;
RX MEDLINE=20500219; PubMed=11048724; DOI=10.1038/35037619;
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
RA Cassell G.H.;
RT "The complete sequence of the mucosal pathogen Ureaplasma
urealyticum.";
RL Nature 407:757-762(2000).
DR EMBL; AE002115; AAF30549.1; -.
DR InterPro: IPR011591; Botulinum.
DR ProDom: PD001963; Botulinum; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 1272 AA; 149378 MW; B110E502B96770F9 CRC64;

Query Match 11.3%; Score 114.5; DB 2; Length 1272;
Best Local Similarity 25.0%; Pred. No. 25;
Matches 52; Conservative 36; Mismatches 69; Indels 51; Gaps 11;

QY 14 IPYITVNGTSQNILSLTFNKQOISYKDIEKNKSVLYFNRGISDIDRLSKQAKYTV 73

```

```

Db 637 IPYSL-----RNLI-SUTLHERVAYIK--HNATNGFVLYNTITSNKDFNLKKNKESY 687
QY 74 HFKN-----GTRKRVLDKAG-----IHTADLINTSDIKAKAISVNVDT----- 109
DQ 688 HLNSLIAHLAIKDEGLKK--DESAGGFLEKLDNDFNLNNFKDEDIVFNVVDLELDPNLIY 745
QY 110 KKQVKQKE-AKANVOQPYTITVNGTSQNILSLTFKNQOQISYKDLENNVKSVLKSNRGI 168
DQ 746 ESQVLDKMRSHVPIPISYKVI-KKQHILEDVEFSLNAYALGSEAYENHIYQOIKS---- 800
QY 169 TDVDLRLSKQAKETVNFKNGTKKVIDLK 196
DQ 801 -----QLKFNV-FLNGSKIKVEVK 818

RESULT 31
Q7PDP5 PRELIMINARY; PRT; 1016 AA.
ID Q7PDP5
AC Q7PDP5;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Erythrocyte membrane protein pPEMP3.
GN Name=PY03960;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=12368865; DOI=10.1038/nature01099;
RA Carlton J.M., Angiueli S.V., Suh B.B., Kooij T.W., Perteau M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallow J.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoals A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
parasite Plasmodium yoelii yoelii.";
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; AABL0100183; EABL5755.1; -.
SQ SEQUENCE 1016 AA; 119602 MW; 1C33F4AE66E30413 CRC64;

Query Match 11.1%; Score 113; DB 2; Length 1016;
Best Local Similarity 22.0%; Pred. No. 25;
Matches 54; Conservative 39; Mismatches 83; Indels 70; Gaps 11;

QY 8 HHGQIPYITVNGTSQNILSLTFN-----KNQOISYKDIEKNKSVLYFNRG--- 57
DQ 422 HNLNKOI-NKITINGEYNELIKS--FNSKTRNTQNNIKQIRHLSDQVDNVLMMQKSDNL 478
QY 58 ---TSDIDLRLSKQA-----KYTVHFK-----NGTKRVVDLKAGIHTADLINT 97
DQ 479 FIKINDYINNLNQIGSFILNNQFLHHNENNNYNNNDNNNGTKETINV--DVHNYKEIQN 536
QY 98 SDI-----KAISVNVDTKKQVK-----DKEAKANVOQPYTITVNGTSQNIL 138
DQ 537 TDISKKNSTFIHKCNIDIKYDKKKKKNVNNQIYDDDNNGKRSLSIQKINEYHNENN 596
QY 139 SNLTFK-----KNQOISYKDLENNKSVLKSNGITDVLRLSKQAKFTVNF 185
DQ 597 NNSTVKKDTIQSAQSVYSDKNENNDYKNLYNDEKFLYLNKLI---ELKTSKMKQLNI 653
QY 186 KNGTKK 191

```


Db	348 S	---LTFDNANININELSKI 366
RESULT 35		
Q9ZHL0		
ID	Q9ZHL0	PRELIMINARY; PRT; 4919 AA.
AC	Q9ZHL0, Q7BY44;	
DT	01-MAY-1999 (TrEMBLrel. 10, Created)	
DT	01-MAY-1999 (TrEMBLrel. 10, Last sequence update)	
DT	25-OCT-2004 (TrEMBLrel. 28, Last annotation update)	
DE	Large supernatant protein 2.	
DE	Name=lep2; OrderedlocusNames=HD1156;	
GN	Haemophilus ducreyi.	
OS	Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;	
OC	Pasteurellaceae; Haemophilus.	
OC	NCBI_TaxID=730;	
OX	[1]	
RN	SEQUENCE FROM N.A.	
RP	STRAIN=35000;	
RC	MEDLINE=39030326; PubMed=9811662;	
RA	Ward C.K., Lumbley S.R., Latimer J.L., Cope L.D., Hansen E.J.;	
RA	RT "Haemophilus ducreyi secretes a filamentous hemagglutinin-like protein";	
RT	J. Bacteriol. 180:6013-6022(1998).	
RL	[2]	
RN	SEQUENCE FROM N.A.	
RP	STRAIN=35000HP / ATCC 700724;	
RC	Munson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,	
RA	Johnson L., Nguyen D., Wang J., Forst C., Hood L.;	
RA	RT "the complete genome sequence of Haemophilus ducreyi";	
RL	Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AF057696; AAC79761.1; -	
DR	EMBL; AE017154; AAP96012.1; -	
DR	PIR; T31105; T31105.	
DR	GO; GO:0004197; P:cysteine-type endopeptidase activity; IEA.	
DR	GO; GO:0009405; P:pathogenesis; IEA.	
DR	InterPro; IPR008619; F11 haemagg.	
DR	InterPro; IPR008638; Haemagg_act_N.	
DR	InterPro; IPR003951; Peptidase_C58.	
DR	InterPro; IPR006473; Peptidase_C58_yt.	
DR	Pfam; PF05594; F11 haemagg; 10.	
DR	Pfam; PF05860; Haemagg_act; 1.	
DR	Pfam; PF03543; Peptidase_C58; 1.	
DR	TIGRFAMs; TIGR01901; adhes_NPXG; 1.	
DR	TIGRFAMs; TIGR01586; yopt_Cys_prot; 1.	
KW	Complete proteome.	
SC	SEQUENCE 4919 AA; 542602 MW; 5779201455CA69A0 CRC64;	
Query Match	11.0%; Score 112; DB 2; Length 4919;	
Best Local Similarity	25.7%; Pred. No. 1.7e+02;	
Matches	52; Conservative 29; Mismatches 67; Indels 54; Gaps 10;	
Qy	18 ITVNGTSQ-----NILSLTFNNKQIQSYKDIE--NKVKSUVLYNFRGISDIDL---RL 65	
Db	577 VTLNDAKLSANKLNDVTNVTNLSKSTLSAGSLTFKVKQNVTLNN---DSELAANNL 632	
Qy	66 SKQAKYTVHFKNCTKRVVDLKGHTADLNTSDIKAISVNDPKQVKQKDEAKANVQP 125	
Db	633 SLNASHNVTLNNKSK-----USA-----QKADIKAVNLTLNDDTTLTAKNLNDIN---S 677	
Qy	126 YTTIVNGTSQNILSLNLTFFK-----NQIQSYKDLNNVKSVLKSNRGITD 170	
Db	678 TTTINNGTIAGIPANITTEKLNNKEKALILAEQNLNFTVNGSHYENKGDIVSKDKATVT-- 736	
Qy	171 VDLRLSKQAKPTVNFKNQTKKV 192	
Db	737 ----FSKNSDPT-----SNGSKLV 751	
RESULT 36		
Q65HY4		
ID	Q65HY4	PRELIMINARY; PRT; 483 AA.

```

RP SEQUENCE FROM N.A.
RX MEDLINE=22255705; PubMed=1236864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlson J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Perteau M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Nature 419:498-511(2002).
RL EMBL; AE014833; AAN35462.1; -.
KW Hypothetical protein.
SQ SEQUENCE 1811 AA; 216655 MW; 8A25116576D5FED1 CRC64;

Query Match 10.9%; Score 111; DB 2; Length 1811;
Best Local Similarity 23.4%; Pred. No. 65;
Matches 60; Conservative 34; Mismatches 74; Indels 88; Gaps 14;

QY 6 HHHHGSQIPYITVN-----GTSQNILSS-----LTFKKNQIISYKDI 44
Db 1404 HNHVDGNIILTWNNINISLTLGSKNPFVSPFKTHTKIKPEPILFNMNQ---KDV 1460

QY 45 ENK-----KSVLYFNRGISDIDLRLS---KQAKYIVH-----FKNG-----TK 80
Db 1461 QNKNQVDYIDILFSTGFCDFISVTCGIYRKKYTCYNISTIPWFGYKNWGLNCKMK 1520

QY 81 RVVDLKAGIHTADLINTSDIKAISVNVDTKKVQKDKAKANVQVPYITVNGTSONILSN 140
Db 1521 DMVSIKIDH--DIKEKNEIKETLNLNLSQIKDK--KNTYK-----N 1559

QY 141 LTFKKNQIISYKDLNNV-----KSVLSKNRGITVDVLRSLKQAKFTVN 184
Db 1560 VNLSHDHQNDYKD-SNNVFNQDKNLYQHNIAYNKEIQKN--IYIDDNKKKKESNN 1616

QY 185 FKNGTKKKVIDLKAGIY 200
Db 1617 MSIKKNLYEKKNVY 1632

RESULT 38
Q81581 PRELIMINARY; PRT; 3896 AA.
AC Q81581
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Hypothetical protein.
GN ORFNames=PFL1645;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]

SEQUENCE FROM N.A.
RP MEDLINE=22255705; PubMed=1236864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlson J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Perteau M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Nature 419:498-511(2002).
KW Hypothetical protein.
SQ SEQUENCE 1811 AA; 216655 MW; 8A25116576D5FED1 CRC64;

```

```

RA Nakao B., Rowley D., Tamaki T., Wang F., Davis R.W.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE014849; AAN36415.1; -.
DR HSSP; Q92831; 1N72.
DR InterPro; IPR001487; Bromodomain.
DR InterPro; IPR000875; Cereoplin.
DR Pfam; PF00439; Bromodomain; 1.
DR PRINTS; PR00503; Bromodomain; 1.
DR SMART; SM00297; BRMO; 1.
DR PROSITE; PS00014; BROMODOMAIN 2; 1.
DR PROSITE; PS00268; CECROPIN; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 3896 AA; 466013 MW; 12B85DF96B07B29C CRC64;

Query Match 10.9%; Score 111; DB 2; Length 3896;
Best Local Similarity 25.7%; Pred. No. 1.6e+02;
Matches 45; Conservative 32; Mismatches 68; Indels 30; Gaps 7;

QY 34 KKNQIISYKDIENKV-KSVLYFNRGISDIDLRLSKQAKYTVHFKNGTKRVDL----- 85
Db 3376 KKKKIKNDNDNNINYKELYENMNSTNENLALKKQKKQKNKTKKKKKOMDILYNDKN 3435

QY 86 KAGIHTADLINTSDIKAISVNVDTKK-----QVXDKKAKANVQVPYITVNGTSONILSN 140
Db 3436 KKGIN-KKFVNTN--YSININKNKRPHNIISYKNKDNKNVQ-----KNNIKKK 3482

QY 141 LTFKKNQIISYKDLNNVSVLKNRGIT---DVDLRLSKQAKFTVNFKNGTKK 191
Db 3483 KTHINKSIVKKNKTNSEISEIKGHEFHQIIEFSDSKKKKKKKKKQNVDESSDNKK 3537

RESULT 39
Q8EWV7 PRELIMINARY; PRT; 240 AA.
AC Q8EWV7
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Hypothetical protein MYPE920.
GN OrderedLocNames=MYPE920;
OS Mycoplasma penetrans.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=28227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=HP-2;
RX MEDLINE=22354719; PubMed=12466555; DOI=10.1093/nar/gkf667;
RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K.,
RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;
RT "The complete genomic sequence of Mycoplasma penetrans, an
RT intracellular bacterial pathogen in humans.";
RL Nucleic Acids Res. 30:5293-5300(2002).
DR EMBL; AP004170; BAC43883.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 240 AA; 27869 MW; 9C98FED493FD003C CRC64;

Query Match 10.9%; Score 110.5; DB 2; Length 240;
Best Local Similarity 24.6%; Pred. No. 7;
Matches 51; Conservative 41; Mismatches 66; Indels 49; Gaps 10;

QY 26 NILSSLTFNKNQOI-----SYKDIENKVS---VLY-----FNRGISDI- 61
Db 8 NMSTASFNKIKQILSNIKQIDSFQDKVNNLSSDNAVHDFKFRYKNRYRNVVSDFD 67

QY 62 DLRLSKQAKYTVHFKNGTKRVDLKAGIHTADLI-----NTSDIKAISVNVDTKK 112
Db 68 DNNLTREEKINLFNQKRTIEDIKKQINFVDYIFSESNNETQASKIKTI-----KGV 121

QY 113 VKDKKAKANVQVPYITVNGTSONILSNLTFKKNQ---QISYKDLNNVSVLKNRGIT 169
Db 122 LSNKEYDLNES-----DINSICDAISNDGKQELLEELTHKEVSDSTTQDSLSLILT 176

QY 170 -DVDLRLSKQAK-FTVNFKNGTKKVID 194

```

Db 177 QDFLSKLKKEGKEISTNFEIKSEKND 203

RESULT 40

Q8RBU3 PRELIMINARY; PRT; 398 AA.
AC Q8RBU3;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Predicted transposase.
GN OrderedLocusNames:TTB0715;
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4;
RX MEDLINE=21992816; PubMed=11997336; DOI=10.1101/gr.219302;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of the T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
DR EMBL; AE013039; AM23977.1; -.
DR InterPro; IPR001959; Transposase_2.
DR InterPro; IPR010095; Tepadset_teng_C.
DR Pfam; PF01385; Transposase_2; 1.
DR Pfam; PF07282; Transposase_35; 1.
DR TIGRFAMs; TIGR01766; tpadset_teng_C; 1.
KW Complete proteome.
SQ SEQUENCE 398 AA; 46552 MW; 6268F889A16A05AE CRC64;

Query Match 10.8%; Score 110; DB 2; Length 398;
Best Local Similarity 24.5%; Pred. No. 13;
Matches 47; Conservative 31; Mismatches 78; Indels 36; Gaps 8;
Qy 21 NGTSNILSSLTFFNQOISYKD-----IENKVSPLYFNRG-ISDIDL 64
Db 116 NGRSVAIFT-----NQCKIKDGHLPFKTNLKLKTRITGKLEVRIPKGSVVVIEIV 169
Qy 65 LSKQAKYTHFKNGTKRVVDLKGIIHT-ADLINTSDIKAISVNDTKQVKDKKAKNVQ 123
Db 170 YEKE---VVEYTKPSKRIAGIDLGNNFVTLVNNIGIRPIVINGKVIKSIQYNNKKAE 226
Qy 124 VPYITVNGTSNILSLTFKKNQOISYKDLNNVSKNSNRGITDVLRLSKQAKFTV 183
Db 227 LMSYVGDGRGTS-NRIEKLTLKNNKI--KDLMHKISRFINVWCKQHDID-----TIVI 276
Qy 184 NFKNGTKKVIDL 195
Db 277 GYNSGWKQIEI 288

Search completed: March 16, 2005, 10:39:41
Job time : 177 secs

the Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 16, 2005, 10:36:37 ; Search time 41 Seconds
(without alignments)
469.350 Million cell updates/sec

Title: US-10-041-775-2
Perfect score: 1016
Sequence: 1 MRGSHHHHSGQIPYITV.....FTVNFNGTKKVIDLKAGIY 200

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	318	31.3	121	2 S15765	hypothetical prote
2	210	20.7	144	2 F89865	hypothetical prote
3	146.5	14.4	141	2 H90016	hypothetical prote
4	136	13.4	106	2 H89982	truncated map-w pr
5	114.5	11.3	640	2 S33734	DNA topoisomerase
6	114.5	11.3	1272	2 H82926	conserved hypothet
7	112	11.0	1127	2 T28317	ORF MSV156 hypothe
8	112	11.0	4919	2 T31105	hypothetical prote
9	107	10.5	483	2 A69745	hypothetical prote
10	107	10.5	520	2 E97813	WASP, N-WASP, MENA
11	104.5	10.3	297	2 C70231	hypothetical prote
12	103.5	10.2	609	2 S57378	hypothetical prote
13	103.5	10.2	1778	2 AF1116	intermalin protein
14	103.5	10.2	3216	2 C90538	hypothetical prote
15	103	10.1	688	2 T43783	hypothetical prote
16	102.5	10.1	410	2 AH1484	probable cell surf
17	102.5	10.1	624	2 PC6003	surface membrane p
18	102.5	10.1	4152	2 T31102	filamentous hemag
19	102	10.0	796	2 T43782	hypothetical prote
20	101	9.9	1536	2 A43855	high-molecular-wei
21	100.5	9.9	405	2 AB1461	B. subtilis vabE p
22	100.5	9.9	450	2 T16440	hypothetical prote
23	100	9.8	451	2 AH1293	hypothetical prote
24	100	9.8	533	2 S52702	REF2 protein - yea
25	99.5	9.8	226	2 E69859	conserved hypothet
26	98.5	9.7	888	2 E82885	hypothetical prote
27	98.5	9.7	1176	2 A33856	surface-layer 125K
28	98	9.6	450	2 E69934	conserved hypothet
29	98	9.6	522	2 D72349	conserved hypothet

30	97.5	9.6	244	2 A86873	transcription regu
31	97.5	9.6	1207	2 S27954	leucine-rich prote
32	97	9.5	647	2 H89988	hypothetical prote
33	97	9.5	3890	2 C89921	hypothetical prote
34	96.5	9.5	408	2 AC1098	B. subtilis vabE p
35	96	9.4	295	2 G97827	hypothetical prote
36	96	9.4	690	2 E84945	glycine-tRNA ligase
37	95.5	9.4	377	2 H82882	hypothetical prote
38	95.5	9.4	521	2 E70129	acid-inducible pro
39	95.5	9.4	682	2 F90603	vipe-like (mycopla
40	95	9.4	298	2 C97183	probable nucleotid
41	95	9.4	746	2 T47237	myosin II heavy ch
42	95	9.4	1039	2 H95115	conserved hypothet
43	95	9.4	1039	2 D37985	hypothetical prote
44	95	9.4	1398	2 H71606	hypothetical prote
45	94.5	9.3	645	1 SUBSMP	serine proteinase

ALIGNMENTS

RESULT 1

S15765
hypothetical protein 1 (h1b 5' region) - Staphylococcus aureus (fragment)
C:Species: Staphylococcus aureus
C>Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004
C:Accession: S15765; S04522
R:Projan, S.J.; Kornblum, J.; Kreiswirth, B.; Moghazeh, S.L.; Eisner, W.; Novick, R.P.
Nucleic Acids Res. 17, 3305, 1989
A>Title: Nucleotide sequence: the beta-hemolysin gene of Staphylococcus aureus.
A:Reference number: S15765; MUID:89263748, PMID:2726469
A:Accession: S15765
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-121 <PRO>
A:Cross-references: UNIPROT:P21223; EMBL:X13404; NID:G46586; PIDN:CAA31768.1; PID:G46587

Query Match 31.3%; Score 318; DB 2; Length 121;
Best Local Similarity 76.5%; Pred. No. 5.1e-15;
Matches 62; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY	4	SHHHHHSQIPYITVNGTSQNLSSLTFTNKNQOISYKDIENKVKSVLYFNRGISDIDL 63
Db	41	SUHHGYSKVHPYAITVNGTSQNLSSLTFTNKNQNISYKDLSDRVKSVLKSDRGISDIDL 100
QY	64	RLSKQAKYTVHFNGTKRWVD 84
Db	101	RLSKQAKYTVHFNGTKKVID 121

RESULT 2

F89865
hypothetical protein SA0841 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: F89865
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1235-1240, 2001
A>Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: F89865
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-144 <KUR>
A:Cross-references: UNIPROT:Q99VA9; GB:BA000018; PID:G13700785; PIDN:BAB42081.1; GSPDB:G

Query Match 20.7%; Score 210; DB 2; Length 144;
Best Local Similarity 44.3%; Pred. No. 1.2e-07;

Science 294, 849-852, 2001
A;Authors: Kraft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Makarewicz, C.; Schluster, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, W.; Weiss, R.; Zimmermann, S.
A;Title: Comparative genomics of *Listeria* species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AF116
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1778 <GLA>
A;Cross-references: UNIPROT:Q8YA32; GB:NC_003210; PIDN:CAC98412.1; PID:g16409711; GSFPDB:1G16
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo0333

```

Query Match      10.2%; Score 103.5; DB 2; Length 1778;
Best Local Similarity 20.2%; Pred. No. 42;
Matches 50; Conservative 38; Mismatches 73; Indels 87; Gaps 11;

QY      17  TITVNGTSONILSLTFKNKQOISYKDIEKNKSVLVFNRGISDID----- 62
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1014  TTVNKKDPAPIISAKT-----EITYDKFSKTEAAF-----LDDIDADTNGDSIVTSNFA 1063
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      63  --LRLSKQAKYTVHFKNKGRVVDLKGAIHTADLINTSDIKAISVNVDTKKQ----- 112
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db      1064  TAVNDKAGDYTVTLNSINS---DGVAGTPTAIIVHVEKEKIATISNTAQQYKVAKIN 1120
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      113  ----VKDKEAKAN-----VQVP--YTITVNGTSON----- 136
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db      1121  ETQFLKDQVHASINASPPTAVLESDFETVVKLDVPFGTYVTIITATNEDCGVSAPKEVSIVV 1180
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      137  -----ILSNLTFKKNQOISYKDLENNKSVLK--SNRGIT---DVDLRLSKQAKFTV 183
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db      1181  RKIPAPEITADKEITYPKFDEVSEAFPLNDIHATISDKNVAITSNFSTDVNLKAGDYTV 1240
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      184  NFNKNGTKK 191
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db      1241  TL-NATNE 1247
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 14

C90538
hypothetical protein MYPU_2110 [imported] - Mycoplasma pulmonis (strain UAB CT1P)
C;Species: Mycoplasma pulmonis
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: C90538
R;Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galissou, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A;Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
A;Reference number: A99512; MUID:21267165; PMID:11353084
A;Accession: C90538
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-3216 <KUR>
A;Cross-references: UNIPROT:Q98QZ9; GB:AL445566; PID:gl4089624; PIDN:CAC13384.1; GSPDB:G
A;Experimental source: strain UAB CT1P
C;Genetics:
A;Gene: MYPU_2110
A;Genetic code: SGC3

Query Match	10.2%	Score 103.5;	DB 2;	Length 3216;
Best Local Similarity	25.6%;	Pred. No. 86;		
Matches	50;	Conservative 33;	Mismatches 83;	Indels 29; Gaps 7;

QY	16	YTTVNGTSONILSSITFNKNQOISYKDIENKYSVL-----YFNRGISDIDLRLSKQA	69
	:	:	:
	:	:	:
Dd	201	YEIKV--SSVNVA SLTSVNVLLISRNGIQEKKVTIIIGIIPAYQSKVDQANDVRPLKD	258
	:	:	:
	:	:	:
QY	70	KYTTFHFKNGTKRVVDLKAGIHHTADLTINTSDIKAI--SNVNDTKQVKDKAKANQVPYT	127
	:	:	:
	:	:	:
Dd	259	NLVISYNNGIDSPDLINKV--ASQISLNDPFHEIVPSDNVDYKVWSIKASEQNAQ-SVI	315
	:	:	:
	:	:	:
QY	128	ITVNGTSONILSNTFTKKNQOISYKOLENNVKSVLKSNRGITDVDDLRLSKQAFTYNFKN	187
	:	:	:
	:	:	:

```

Db      316  LTIRKTSKTHGNNAFYKDYVEYDNLSSDMFELYNR-----VEANPSVNFQT 365
Oy      188  GTKK-----VIDLK 196
           |       :   |||
Db      366  QAKRSAYEYTSIDLK 380

RESULT 15
T43783
hypotheical protein 688 [imported] - slime mold (Dictyostellium discoideum) mito
C:Species: mitochondrion Dictyostellium discoideum
C:Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 09-Jul-2004
C:Accession: T43783
R:Ogawa, S.; Yoshino, R.; Angata, K.; Pi, M.; Iwamoto, M.; Kuroe, K.; Matsuo, K.
submitted to the EMBL Data Library, December 1996
A:Description: The mitochondrial DNA of Dictyostellium discoideum. Complete sequ
A:Reference number: Z22666
A:Accession: T43783
A:Status: preliminary; translated from GB/EMBL/DDBB
A:Molecule type: DNA
A:Residues: 1-688 <OGA>
A:Cross-references: UNIPROT:Q34312; EMBL:AB000109; PIDN:BAA78086.1
C:Genetics:
A:Genome: mitochondrion
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 11
C:Keywords: mitochondrion

```

Query Match	10.1%	Score 103;	DB 2;	Length 688;
Best Local Similarity	24.3%;	Pred. No. 14;		
Matches 49;	Conservative 41;	Mismatches 48;	Indels 64;	Gaps 13;
Qy	26	NILSSLTFNKNGQISYKD	IKENKVSVLVFNRG	-----SDIDLRLSKQAKYTVHFNGT 79
Db	518	NVLVSDSNEVQ-----	VTNKINYIY--QGI	INDKFNKIDLYL--PSGHYEFDEFGD 567
Qy	80	KRVVDLKAGIHTADLNT	-----SDIKAISVNV	DTKKQVKQKAKANKANVQVPYITVTNGTSQ 135
Db	568	REV-----	YMNFTGQORSEIEKLS	ISGKNK--IKENSMIGYQLMY----- 605
Qy	136	NILSN--LTFKKNQQ--	--ISYKDLNENNVKS	VLKSNRGI-----TDVDLRLSKQ 178
Db	606	--LNNKEMTKREKQDK	IKLSYREMKKEERK	IKVWKYLTIINNIIENNYMTDINIRLSKN 663
Qy	179	AKFTVNFNGTKKVIDLK	AGIY 200	
Db	664	LMITGOLRK-EKKI--	MEAGIW 682	

RESULT 16
AH1484
probable cell surface protein (LPXTG motif) [Imported] - *Listeria innocua* (strain
C/Species: *Listeria innocua*
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AH1484
R/Glaser, P.; Franguel, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, H.;
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.;
D.; Jones, L.M.; Karst, U.
Science 204, 849-852, 2001
A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournan,
Ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.;
A/Title: Comparative genomics of *Listeria* species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AH1484
A/Status: preliminary
A/Molecule type: DNA
A/Molecule type: DNA
A/Residues: 1-410 <GLA>
A/Cross-references: UNIPROT:Q92EP3; GB:AL592022; PIDN:CAC95648.1; PID:gl6412844
A/Experimental source: strain Clip11262
C/Genetics:
A/Gene: lin0415

Query Match	10.1%;	Score 102.5;	DB 2;	Length 410;
Best Local Similarity	21.7%;	Pred. No. 8.1;		

Matches	57; Conservative	35; Mismatches	74; Indels	97; Gaps	13;
QY	23	TSQNILSSITFN-----	KNQOISYKDIEKNKSVLYFNRGISDIDLRLSKQAK	70	
Db	9	TTVALLSMVVFIPISMWADAASSTSESNLYTKDVRNG-----	FYFVGVEDVQLEKGSYK	63	
QY	71	YTVHFK-----	NGTKRW--DLKAGIHTADLINTSDIKAISVNV-DTKKQVXD-----	115	
Db	64	YTVAYEANDMQATNTVTQGSAKAGLF--	IPNSSGVELDSNNVTITQNNVVVDVANDGNK	120	
QY	116	-----	KBKANKV-----QVPYT-----	IT-----	VNGT 133
Db	121	VFTHIFEFTVKEDTKADIGTYLGAGSVIPTPDATIINKNVSVTNETPAEQAEAPVINA	180		
QY	134	SNILSNLFFKKNQOISYK-----	DLENNVK-----SVLKSNGITDVLRLSKQAKFT--	182	
Db	181	DKTIEQNETFDALNEVTAQKGDLTQSIKVTNTVDTTKSGEYDVDSYVNSSKLTIT	240		
QY	183	-----	VNFKNKGTKKVIDLK	196	
Db	241	KSVKVTVPVETKNTAPVIDAK	263		
RESULT 17					
PC6003					
surface membrane protein lmp4 - Mycoplasma hominis (fragment)					
N/Alternate names: hypothetical 624 protein; lmp4 protein					
C/Species: Mycoplasma hominis					
C/Date: 15-Aug-1996 #sequence_revision 18-Oct-1996 #text_change 09-Jul-2004					
C/Accession: PC6003					
R/Ladefoged, S.A.; Jensen, L.T.; Brock, B.; Birkelund, S.; Christiansen, G.					
J. Bacteriol. 178, 2775-2784, 1996					
A/Title: Analysis of 0.5-kilobase-pair repeats in the Mycoplasma hominis lmp gene system					
A/Reference number: JC6009; MUID:96213016; PMID:8631664					
A/Accession: PC6003					
A/Molecule type: DNA					
A/Residues: 1-624 <LAD>					
A/Cross-references: UNIPROT:Q49548; EMBL:X95601; NID:g1197335; PIDN:CAA64859.1; PID:g119					
C/Genetics:					
A/Gene: lmp4					
A/Genetic code: SGC3					
C/Superfamily: surface-located membrane protein lmp3; tetratricopeptide repeat homology					
C/Keywords: duplication; membrane protein					
Query Match					
Best Local Similarity					
Matches					
QY	18	ITVNGTS--QNILSSIL-----	TFNKNQOISYKDIEKNKSVLYFNRGISD---	60	
Db	307	VASNNATASQSAKSLDAKVTEITKKLETFNKDKDAKFELQTRKDIDEFIKQIENDPQ	366		
QY	61	-----	IDLRLSKQAKYTVHFKNGTKRVVD-----	LKAGI-----	89
Db	367	TKKNYQNIKNLKDCKAEKNSITFSNNKKEIQDANKSLQDELNNAKITKKGITDFYNSKK	426		
QY	90	HTADLINTSDIKAISVN-----	VDTKQVKDEKANKVQVPY-TITVNGTSQNILSNLTF	143	
Db	427	QLELDITDADKVGTEADTILDHYKNISD--	ASKNEBIQATQKINDIKKIETKIQE	484	
QY	144	KKNQOIS-YKDLENNKSVLKSNGITDV--	DLRLSKQAKFT-VNFKNGTKKVIDLK	196	
Db	485	KKNEFSQFEQIKNELQSPI-NKDLKDQKYSIRTKIENKINGVSSINKNSKIQDIE	540		
RESULT 18					
T31102					
filamentous hemagglutinin 1 - Haemophilus ducreyi					
C/Species: Haemophilus ducreyi					
C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004					
C/Accession: T31102					
R/Ward, C.K.; Lumbley, S.R.; Latimer, J.L.; Cope, L.D.; Hansen, E.J.					
J. Bacteriol. 180, 6013-6022, 1998					

A/Title: Haemophilus ducreyi secretes a filamentous hemagglutinin-like protein.					
A/Reference number: Z20984; MUID:99030326; PMID:9811662					
A/Accession: T31102					
A/Status: preliminary; translated from GB/EMBL/DBJ					
A/Molecule type: DNA					
A/Residues: 1-4152 <WAR>					
A/Cross-references: UNIPROT:Q9ZHL3; EMBL:AF057695; NID:g3929017; PID:g3929018; PIDN:AAC7					
C/Genetics:					
A/Gene: lspA1					
Query Match					
Best Local Similarity					
Matches					
QY	10	HGSPYPTITVNGT-----	SON-----	ILSSITFNKNQOISYKDIEKNVK	49
Db	341	YGSNIKFVVTDKGAGVKHKGIIIFSENDINIKMDGNSLKELYAKKDIILAKDLELTEK	400		
QY	50	SVLYFNRG-----	SDIDLR-LSQAKYTVHFKNGTKRVVDLKAGIHTADLINTSDIKAI	103	
Db	401	QQLQANNKIILNSTGKINLRNASEVSADNVNVKSENLALENASMSANSLDIVITK---	I	456	
QY	104	SVNVDTKKQVKDEKANKVQVPYITVNGTS-----	QNILSNLTFKNQOISYKDLKLE-	155	
Db	457	EVRSSKVSAGTANIKAS-----	NITLDGSSVAVANKITLVNTNNATLNNQSKLSAKDMEL	511	
QY	156	-----	NNVKSVLKSNGRGTVDVLRLSKQA-----	KFTVN-----	FKNGTKKVIDLKA 197
Db	512	NVTNITLNNYSKUSAQKANKITENLTNGEASLVAEKLIDINAIDKITNNGT--	IAGLTA	569	
QY	198	GI 199			
Db	570	NI 571			
RESULT 19					
T43782					
hypothetical protein 796 [imported] - slime mold (Dictyostelium discoideum) mitochondrio					
C/Species: mitochondrion Dictyostelium discoideum					
C/Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 09-Jul-2004					
C/Accession: T43782					
R/Ogawa, S.; Yoshino, R.; Angata, K.; Pi, M.; Iwamoto, M.; Kuroe, K.; Matsuo, K.; Morio,					
submitted to the EMBL Data Library, December 1996					
A/Description: The mitochondrial DNA of Dictyostelium discoideum. Complete sequence, gen					
A/Reference number: Z22666					
A/Accession: T43782					
A/Status: preliminary; translated from GB/EMBL/DBJ					
A/Molecule type: DNA					
A/Residues: 1-796 <OGA>					
A/Cross-references: UNIPROT:Q9TGM3; EMBL:AB000109; PIDN:BAA78085.1					
C/Genetics:					
A/Genome: mitochondrion					
C/Keywords: mitochondrion					
Query Match					
Best Local Similarity					
Matches					
QY	16	YTTIT--VNGTSQNTLSLTFNKNQOISYKDIEKNKSVLYFNRGISDIDLRLSKO--	AKY	71	
Db	15	YTTTGWVKSTHGINIISFNRTIQYQRISSNKNK-----	NRNLSRVLNLKEKGTARV	69	
QY	72	TVHFKNGTK--RVVDLKAGIHTADLINTSDIKAISVNVDTKKQVKDEKANKVQVPYIT	129		
Db	70	TKLNKRNKGRVQVK-----	DILN-----	AKLVQTPIVV-	101
QY	130	VNGTSQNTLSLTFNKNQOISYKDLENNKSVLKSNGRGTDV--	DLRLSKQAKFTVNFK	186	
Db	102	-----	BEIKNISFA--SQMVHKE-HDNLPTLRSKQRMLLKAYFDQLRIN-NLKAVNNIF	151	
QY	187	NGTKKVIDLKA	197		
Db	152	TLKKRVNLKS	162		

Matches 47; Conservative 32; Mismatches 58; Indels 48; Gaps 11;

Qy 17 TITVNG-----TSQNISSLTFNKNQOISYKDIEN-----KVKSVLYFNRGISDIDLRL 65

Db 77 TIASNGDLLETTNEELKELASSDNR---YKDIKAYRLREAKKVSEY---INTKYRY 130

Qy 66 -----SKQAKYTVHFKNKGRVVDLKGAIHTADLINTSDIKAISVNVVDIKKQVKDEAK 119

Db 131 PTTTIQKFFANVEI---TSGERSVQTEKTPVQIBETTN-NDIK---VSLNTSYSSDDAEVA 184

Qy 120 A---NVQVPYTTVNGTSONILSN-----LTFKNQOISYKDLNNVKSVLK 163

Db 185 ADIFNSQIPETTTIDNTKNLITKNKNVTIYPKTFYDFTLQNTWLVPKEQE---EAVIK 241

Qy 164 SNRGI 168

Db 242 EPIGI 246

RESULT 24

S52702

REF2 protein - yeast (Saccharomyces cerevisiae)

N/Alternate names: protein YD9346.06; protein YDR195W

C/Species: Saccharomyces cerevisiae

C/Date: 19-May-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004

C/Accession: S52702

R/Oliver, K.; Harris, D.

submitted to the EMBL Data Library, March 1995

A/Reference number: S52697

A/Accession: S52702

A/Molecule type: DNA

A/Residues: 1-533 <COLI>

A/Cross-references: UNIPROT:P42073; EMBL:Z48784; NID:g755782; PID:g755788; MIPS:YDR195W

A/Experimental source: strain AB972

C/Genetics:

A/Genes: SGD:REF2

A/Cross-references: SGD:S0002603; MIPS:YDR195W

A/Map position: 4R

Query Match 9.8%; Score 100; DB 2; Length 533;

Best Local Similarity 25.5%; Pred. No. 16;

Matches 49; Conservative 36; Mismatches 71; Indels 36; Gaps 7;

Qy 19 TVNGTSONILSLTFNKNQOISYKDIENKVKSVLYFNRGISDIDLRLSKQAKYTVHFKNK 78

Db 247 TASTSSMKKLKGLFNKNEAKSTESLPTSSKKLSFSKYNKDDADMKL-----G 297

Qy 79 TKRVVDL-----KAGIHTADLINTSDIKAISVNVDTKKQVKD---KEAKANVQVPYTTIT 129

Db 298 TKRSIDVDFKVNPEASTVASNIISSTSGSSTTTVATPASSEPLKTKTSVQ----- 351

Qy 130 VNGTSONILSNLTFNKNQOISYKDLNNVKSVLK-----SNRGITVDLRLSKQAKFTV 183

Db 352 -DSNVQSILNRGPKKARISSIKFLDSS--QLIKVYGDLPNQGLQVSPQLKKILK--- 405

Qy 184 NFKNGTKKVIDL 195

Db 406 PFKEGEPKEIIL 417

RESULT 25

E69859

conserved hypothetical protein ykoI - Bacillus subtilis

C/Species: Bacillus subtilis

C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C/Accession: E69859

R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berton

C.; Bron, S.; Brouillette, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabbet, C.; Ferrari, E.

Nature 390, 249-256, 1997

A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gall

iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holbappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,

A/Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seror

akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K

A/Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A/Reference number: A69580; MUID:98044033; PMID:9384377

A/Accession: E69859

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Residues: 1-226 <KUN>

A/Cross-references: UNIPROT:O34551; GB:Z99110; GB:AL009126; NID:g2633472; PIDN:CAB13184.

A/Experimental source: strain 168

C/Genetics:

A/Genes: ykoI

Query Match 9.8%; Score 99.5; DB 2; Length 226;

Best Local Similarity 23.4%; Pred. No. 6.3;

Matches 50; Conservative 32; Mismatches 69; Indels 63; Gaps 9;

Qy 18 ITVNGTSONILSLTFNKNQOISYKDIENKVKSVLYFNRGISDIDLRLSKQAK-YTVHF 76

Db 24 LIIRQTHENVLSKETVVVKVEASY---EGKVTKATQ-----SKDKTYDITL 68

Qy 77 NGTKRVVDLKGAIHTADLINTSDIKAISVNVDTKKQVKDEAKANVQVPYTTITVNGTSQN 136

Db 69 N-PKGYTFVKADALSADILSNRVKAVNPSAMTEKEA---EHLALERVPGTVK-KQTRQS 123

Qy 137 ILSNLTFFKN-----QQISYKD-----LENNVKSVLKSNRGIT-- 169

Db 124 QVATYTTIQKDGKTYEVKVDQAKTVLSADQISKQOQKTPITKKEAKTIAERTGTGAD 183

Qy 170 -----DVLRLSKQAKFTVNFKNK 188

Db 184 DADLESEGLTFEVDVLDPNKEATVKINAYTG 217

RESULT 26

E82885

hypothetical protein UU481 [imported] - Ureaplasma urealyticum

C/Species: Ureaplasma urealyticum

C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C/Accession: E82885

R/Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.

submitted to GenBank, February 2000

A/Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min

A/Reference number: A82870

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-888 <GLA>

A/Cross-references: GB:AB002145; GB:AF222894; NID:g6899476; PIDN:AAF30893.1; GSPDB:GN001

A/Experimental source: serovar 3; biovar 1

C/Genetics:

A/Genes: UU481

A/Genetic code: SGC3

Query Match 9.7%; Score 98.5; DB 2; Length 888;

Best Local Similarity 21.3%; Pred. No. 39;

Matches 54; Conservative 42; Mismatches 86; Indels 71; Gaps 13;

Qy 8 HHGSGQIPYTTITVNGTSONILSLTFNKNQOISYKDIEN-----KVKSVLYFN---RG 57

Db 547 YNDGDRLVLEYRNNSSTKKEVSLTLNKNQEDYSVNIENLTYNREYTLKNIKINHNTNN 606

Qy 58 ISDIDLRLSKQAKYTV-----HFKNKT-----KRVVDLKGAIHTAD----- 93

Db 607 ITSIDISKLNDKTFVVEESVTKLTFNFKNITNDEKQKQWSRIMEFNI---TNDDEWFQS 663

Qy 94 LINTSDIKAISVNVDTKK-----QVKDEAKANVQVPYTTITVNGTSONILSNLTF--- 143

A:Cross-references: UNIPROT:Q99UE3; GB:BA000018; PID:g13701233; PIDN:BAB42528.1; GSPDB:G
A:Experimental source: strain N315
C:Genetics:
A:Gene: ebhB

Query Match 9.5%; Score 97; DB 2; Length 3890;
Best Local Similarity 21.4%; Pred. No. 3e+02;
Matches 49; Conservative 31; Mismatches 73; Indels 76; Gaps 9;

QY 17 TITVNGTSQNLSSLTFFNKKQISYKDIENKVKSVLYFNRG--ISDIDLRLSKQAKYTVH 74
DB 1036 TAVSNDEKQSIITAP-MNKNQMI-----RGVLASTDPVVDNNGNVTLH 1078
QY 75 FNGTKRVVD-----LKAGHTADLINTSDI---KALSUNVDTKKQVKDEKAKAN 121
DB 1079 YRDGSSSTLDTATNVMTYEPVVKSEYQTANAATATVTIAKGOSFNTGDIKQY----- 1130
QY 122 VQVPVYIT-----VNGTSQNLSSLTFFNKKQISYKDIENKVKSVLKSNNR----- 166
DB 1131 ----FTLSNGQALPNGTFTNITSDRPTIPTAQEVSQNNAGTQLYHIVASNAHYHKTEDFYI 1186
QY 167 --GITDV-----DLRLSKQAKFTVNFNKGTKKVIDLKA 198
DB 1187 SLKIVDVKQPEGQRVYRTSTYDLTDTDEISKVKQAFINARDVITLAE 1235

RESULT 34
AC1098
B. subtilis YabE protein homolog lmo0186 [imported] - Listeria monocytogenes (strain EGD
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AC1098
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussauget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kref, J.; Kuhn, M.; Kunat, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Sincos, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AC1098
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-408 <GLA>
A:Cross-references: UNIPROT:Q8YAB4; GB:NC_003210; PIDN:CAC98401.1; PID:g16409543; GSPDB:
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo0186

Query Match 9.5%; Score 96.5; DB 2; Length 408;
Best Local Similarity 24.2%; Pred. No. 21;
Matches 50; Conservative 34; Mismatches 74; Indels 49; Gaps 10;

QY 11 GSOIIPY-----TITVNGTSQNLSS-----LTENKN-----QQISYKDIENKVKSVL- 52
DB 89 GMEIKYLPARQITINDNGTKDQVSWTKYANVALLDKNITTPQDVNLVALDTKLNGLUE 148
QY 53 -YFNRGISDIDLRLSKQAKYTVHFNKGTRVVDLKAHTADLINTSDIKAISVNVDTKK 111
DB 149 VNINRAI-----QLSLQ-----NGAKDVTWTTTKTKVSDLLAEKNIK-----LDQDD 190
QY 112 QV---KDEAKANVQVPYITVNGTSQNLSSLTFFNKKQISYKDIENKVKSVLKSNNR 168
DB 191 RVSPAKDSNLKERMTEVTVY-VNSKAE-----KKNEQVKFETVYKEDDSDSLNKGEKV 241
QY 169 TDVDLRLSKQAKFTVNFNKGTKKVIDL 195
DB 242 VOEGKNGKKVVEYKVTFFENGKKKRDV 268

RESULT 35
G97827
hypothetical protein RC1023 [imported] - Rickettsia conorii (strain Malish 7)

C:Species: Rickettsia conorii
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: G97827
R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ri
Science 293, 2093-2098, 2001
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A:Reference number: A97700; MUID:21442074; PMID:11557893
A:Accession: G97827
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-295 <KUR>
A:Cross-references: UNIPROT:Q92GU9; GB:AE006914; PIDN:AAL03561.1; PID:g15620140; GSPDB:
C:Genetics:
A:Gene: RC1023

Query Match 9.4%; Score 96; DB 2; Length 295;
Best Local Similarity 23.3%; Pred. No. 15;
Matches 49; Conservative 36; Mismatches 65; Indels 60; Gaps 10;

QY 10 HGSQIIPY--TITVNGTS-----QNILSSLTFFNKKQISYKDIENKVKSVLYFN 56
DB 105 HGSLLQYLKRSVNPDSFISQIKSELIMNLLSSL--SRSVQVSNKEID-----VAILSS 157
QY 57 GISDIDLRL-----SKQAKYTVHFNKGTRVVDLKAHTADLINTSDIKAISVNV 107
DB 158 DQKVEISMQVFTSKDGGNKAPTQNNLNKRLKCADVK-----KSLYDNF 203
QY 108 DTKQVKQKDEAKANVQVPYITVNGTSQNLSSLTFFNKKQISYKDIENKVKSVLKSNNR 167
DB 204 ATMQIITDKLSK-----IEGVKQTVIKDLTPDKASNVF--EVNNKFEITLVCSKK 251
QY 168 ITVDLRLSKQAKFTVNFNKGTKKVIDLKA 197
DB 252 ILNVN---EDENNYVNVFL--TNKKISQKA 276

RESULT 36
E84945
glycine-tRNA ligase (EC 6.1.1.14) beta chain [imported] - Buchnera sp. (strain APS)
N:Alternate names: glycyl-tRNA synthetase beta chain
C:Species: Buchnera sp.
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 03-Jun-2002
C:Accession: E84945
R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp.
A:Reference number: A84930; MUID:20445173; PMID:10993077
A:Accession: E84945
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-690 <STO>
A:Cross-references: GB:AP000398; GSPDB:GN00144
A:Experimental source: strain APS
C:Genetics:
A:Gene: glys; BU135
C:Superfamily: glycine-tRNA ligase beta chain
C:Keywords: ligase

Query Match 9.4%; Score 96; DB 2; Length 690;
Best Local Similarity 21.7%; Pred. No. 42;
Matches 55; Conservative 35; Mismatches 54; Indels 110; Gaps 12;

QY 22 QTSQ---NILS--SLTFNKN-----QQISYKDIENKVKSVLYFN 63
DB 11 GTEELPARLLSKISLYFYKFNFKELDFHNSYKNIK-----YFSTPRRLALKIKDIDI 63
QY 64 RLKSKQAKYTVHFNKGTRVVDLKAHTADLINTSD-----IKAISVNVV---- 108
DB 64 -----TERFVEIKK--RGFSINSYDKDGFLEATAARWLKHGGINQAI 106
QY 109 -----TKKQVKDKE-----AKANVQVPYTI----- 128
DB 107 RLKNEKGWLFYKTRKKGNIESLIPKITESALKNISIKKSMRWQDNQKFSRPIRNVI 166

QY 129 -----TVNGTSQNLSTFKKQKQISYKDLNNVSVLKSNGRTDVLRLSKQ 178
| | | | | : | | | | : | | | | : | | | | : | | | | :
Db 167 LLDKKVIPGDVFNITSKNLLQHLSSKQSIKIDAKDYPKILLERKNIADYFIR---K 223
| | | | | : | | | | : | | | | : | | | | : | | | | :
QY 179 AKFTVNFKNQTKKV 192
| | | | | : | | | | : | | | | : | | | | : | | | | :
Db 224 EKIIEDIENIAKKI 237
| | | | | : | | | | : | | | | : | | | | : | | | | :
RESULT 37
H82882
hypothetical protein UU497 [imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: H82882
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mir
A:Reference number: A82870
A:Accession: H82882
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-377 <GLA>
A:Cross-references: GB:AE002148; GB:AF222894; NID:G6899495; PIDN:AAF30909.1; GSPDB:GN001
A:Experimental source: serovar 3; biovar 1
C:Genetics:
A:Gene: UU497
A:Genetic code: SGC3

Query Match 9.4%; Score 95.5; DB 2; Length 377;
Best Local Similarity 24.3%; Pred. No. 22;
Matches 54; Conservative 30; Mismatches 75; Indels 63; Gaps 12;
QY 22 GTSTNLSTLTNKN-QQISYKDLENKSVLY-----FNRGSDIDLRLSKQAKYTVHFK 76
| | | | | : | | | | : | | | | : | | | | : | | | | :
Db 85 GNKINVDINPEYKNLQMLSPK-LENLKPNTTYKITKFNITGHEVDLTQIKDSLFTTKAE 143
| | | | | : | | | | : | | | | : | | | | : | | | | :
QY 77 NGTKRVVDLKAGIHTADLINTSDIKAI SVNVTTKQVKDKAKANVQVYTTIVNGTSON 136
| | | | | : | | | | : | | | | : | | | | : | | | | :
Db 144 NLLPDIPNLPNSVG-----TEIKDIKTSLSDNA-----TKVNVNVNLEIN-QSTLEN 190
| | | | | : | | | | : | | | | : | | | | : | | | | :
QY 137 ILSNLTFFKNQKQISYKDLNNVK-SVLKSNRGITD-----VDLRLSK 177
| | | | | : | | | | : | | | | : | | | | : | | | | :
Db 191 QVARLYKSDN-SWK-LSNTLKIAEKNSFVLDGLISNRKYLKELIIGSQSDNLNTN 248
| | | | | : | | | | : | | | | : | | | | : | | | | :
QY 178 -QAKFTVN-----FKNGTKKVIDLK 196
| | | | | : | | | | : | | | | : | | | | : | | | | :
Db 249 AQTKITTNLSLPTTAPKPEIKSVLIDTKFDNHPSSLNLK 290
| | | | | : | | | | : | | | | : | | | | : | | | | :

RESULT 38
E70129
acid-inducible protein (act206) homolog - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
C:Accession: E70129
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:9805943; PMID:9403685
A:Accession: E70129
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-521 <KLE>
A:Cross-references: UNIPROT:O51253; GB:AE001134; GB:AE000783; NID:G2688133; PIDN:AAAC6663
A:Experimental source: strain B31
C:Superfamily: apolipoprotein N-acyltransferase

Query Match 9.4%; Score 95.5; DB 2; Length 521;

Best Local Similarity 21.4%; Pred. No. 32;
Matches 52; Conservative 35; Mismatches 77; Indels 79; Gaps 10;
QY 14 IPTTIV-----NGTSQNLSTLT-----KQKQIS 40
| | | | | : | | | | : | | | | : | | | | : | | | | :
Db 91 IPYSLTGLYFLYSLSFKNKTMSITMLFTFYDSRSIGFLAYPWGLAAFTVNNFNLIQ 150
| | | | | : | | | | : | | | | : | | | | : | | | | :
QY 41 YKDLENK--VKSPLYF-NRGISDIDLRLSK-----QAKYTVHFKNGTKRVVDLKAG 88
| | | | | : | | | | : | | | | : | | | | : | | | | :
Db 151 IADIFGVFFSVFVFLNSGIADFLIHKNTNLLNTAFLLITASTTYGMICKIELKNL 210
| | | | | : | | | | : | | | | : | | | | : | | | | :
QY 89 IHTADLINTSDIKAISVNVVD-----TKQVKD-----KEAKNVQVPTTIVNGTSON 136
| | | | | : | | | | : | | | | : | | | | : | | | | :
Db 211 L--AKEIDSLNIAAIQLNTDPLPGNDKKGIRDSIBITEQALKENPKIEFV-----WSEG 264
| | | | | : | | | | : | | | | : | | | | : | | | | :
QY 137 ILSNLTFFKNQKQISYKDLNNVSVLKSNGRTDVLRLSKQAKFTVNFKNQTKKVIDLK 196
| | | | | : | | | | : | | | | : | | | | : | | | | :
Db 265 VLTYPFSKQDHPKSSDLHNLKNFIKEHK-----IPFAIGAPSNLDKA 308
| | | | | : | | | | : | | | | : | | | | : | | | | :
QY 197 AGI 199
| | | | | : | | | | : | | | | : | | | | : | | | | :
Db 309 IGI 311
| | | | | : | | | | : | | | | : | | | | : | | | | :
RESULT 39
F90603
vlpE-like (mycoplasma hyorhinis) lipoprotein [imported] - Mycoplasma pulmonis (strain UA
C:Species: Mycoplasma pulmonis
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C:Accession: F90603
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Sanson, D.; Gallison, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
A:Reference number: A99512; MUID:21267165; PMID:11353084
A:Accession: F90603
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-682 <KUR>
A:Cross-references: UNIPROT:Q98P18; GB:AL445566; PID:gl4090149; PIDN:CAC13907.1; GSPDB:G
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPU 7340
A:Genetic code: SGC3

Query Match 9.4%; Score 95.5; DB 2; Length 682;
Best Local Similarity 20.5%; Pred. No. 45;
Matches 54; Conservative 39; Mismatches 55; Indels 115; Gaps 12;
QY 16 YTTIVNGT-----SQNTLSLTFN-----KNQKQISYKDIENKSVLYFNRG-- 57
| | | | | : | | | | : | | | | : | | | | : | | | | :
Db 294 YTPTLDGSDYKSFEEYNRIKEILTSLNFSVKBEKHILSFSEKSS-----DFNLGLL 348
| | | | | : | | | | : | | | | : | | | | : | | | | :
QY 58 -----ISDIDLRLS-----KQAKYTVHFKN----- 77
| | | | | : | | | | : | | | | : | | | | : | | | | :
Db 349 AKEGNDVGFPPKTELESKEISDIDMQITFSPKDDAVCEKNGKPFIEFLAQSLFETIE 408
| | | | | : | | | | : | | | | : | | | | : | | | | :
QY 78 -----GTRKVVLD-----KAGIHTADLINTSDIKAISVNV 107
| | | | | : | | | | : | | | | : | | | | : | | | | :
Db 409 KNIEWKYTAEKIMEDFAKGERVTELENLYNHFKDKANIHPPML-----KDLTTNV 462
| | | | | : | | | | : | | | | : | | | | : | | | | :
QY 108 DTKQVKQKKEAKANVQVYTTIVNGTSON1-----LSNLTFFKNQKQISYKDLNNKSVL 162
| | | | | : | | | | : | | | | : | | | | : | | | | :
Db 463 D-GKQVKIADKNINLKGQVATLKG-SKGVNDYQGSATLIFEVDYRGSKQKQVEITINDFA 520
| | | | | : | | | | : | | | | : | | | | : | | | | :
QY 163 K-----SNRGITDVLRLS 176
| | | | | : | | | | : | | | | : | | | | : | | | | :
Db 521 KVAPHEENVGKNRLQDFDIRAS 543
| | | | | : | | | | : | | | | : | | | | : | | | | :

RESULT 40
C97183
probable nucleotidyltransferase [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum

Search completed: March 16, 2005, 10:43:23
Job time : 45 secs

1116 Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 16, 2005, 10:42:42 ; Search time 137 Seconds
(without alignments)
482.542 Million cell updates/sec

Title: US-10-041-775-2
Perfect score: 1016
Sequence: 1 MRGSHHHHGSQIPYITV.....FTVNFNGTKKVIDLKAGIY 200

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1401741 seqs, 330541175 residues

Total number of hits satisfying chosen parameters: 1401741

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA:*
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/FCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1016	100.0	200	14	US-10-041-775-2
2	626	61.6	131	9	US-09-982-992A-2
3	626	61.6	131	14	US-10-041-775-4
4	440	43.3	150	8	US-08-781-986A-5238
5	440	43.3	150	15	US-10-329-624-5238
6	210	20.7	151	8	US-08-781-986A-5233
7	210	20.7	151	15	US-10-329-624-5233
8	142	14.0	141	17	US-10-470-048B-221
9	133	13.1	343	15	US-10-378-674-4
10	131	12.9	316	15	US-10-378-674-6
11	117.5	11.6	391	14	US-10-057-531A-7
12	117.5	11.6	391	14	US-10-057-532A-7
13	116.5	11.5	512	9	US-09-813-820-6
					Sequence 2, Appli
					Sequence 2, Appli
					Sequence 4, Appli
					Sequence 5238, Ap
					Sequence 5238, Ap
					Sequence 5233, Ap
					Sequence 5233, Ap
					Sequence 221, App
					Sequence 4, Appli
					Sequence 6, Appli
					Sequence 7, Appli
					Sequence 6, Appli

14	115	11.3	374	15	US-10-282-122A-72043	Sequence 72043, A
15	114.5	11.3	393	14	US-10-057-531A-3	Sequence 3, Appli
16	114.5	11.3	393	14	US-10-057-532A-3	Sequence 3, Appli
17	114	11.2	211	9	US-09-813-820-4	Sequence 4, Appli
18	113.5	11.2	246	15	US-10-282-122A-53616	Sequence 53616, A
19	111	10.9	963	15	US-10-282-122A-61097	Sequence 61097, A
20	110	10.8	1220	14	US-10-193-764-28	Sequence 28, Appli
21	110	10.8	1226	14	US-10-193-764-26	Sequence 26, Appli
22	108.5	10.7	371	15	US-10-404-667-5	Sequence 5, Appli
23	106	10.4	345	9	US-09-813-820-7	Sequence 7, Appli
24	105	10.3	560	15	US-10-378-674-2	Sequence 2, Appli
25	104	10.2	159	9	US-09-813-820-2	Sequence 2, Appli
26	104	10.2	1051	15	US-10-282-122A-70558	Sequence 70558, A
27	103.5	10.2	609	15	US-10-369-493-2021	Sequence 2021, Ap
28	103.5	10.2	1778	15	US-10-282-122A-60961	Sequence 60961, A
29	102.5	10.1	184	10	US-09-345-373-30	Sequence 30, Appli
30	102.5	10.1	184	14	US-10-075-446-30	Sequence 30, Appli
31	102.5	10.1	184	14	US-10-035-212-30	Sequence 30, Appli
32	102.5	10.1	184	17	US-10-901-210-30	Sequence 30, Appli
33	101.5	10.0	371	15	US-10-404-667-2	Sequence 2, Appli
34	101.5	10.0	775	16	US-10-437-963-179404	Sequence 179404, A
35	101.5	10.0	3241	9	US-09-841-786-1	Sequence 1, Appli
36	101.5	10.0	3241	15	US-10-647-057-1	Sequence 1, Appli
37	101	9.9	1095	14	US-10-193-764-65	Sequence 65, Appli
38	101	9.9	1536	13	US-10-032-880-2	Sequence 2, Appli
39	101	9.9	1536	14	US-10-193-764-63	Sequence 63, Appli
40	101	9.9	1536	17	US-10-681-171-2	Sequence 2, Appli
41	100	9.8	262	9	US-09-895-828-457	Sequence 457, App
42	100	9.8	262	14	US-10-114-666-457	Sequence 457, App
43	100	9.8	533	16	US-10-477-369-47	Sequence 47, Appli
44	99.5	9.8	177	17	US-10-842-989-27	Sequence 27, Appli
45	99.5	9.8	330	14	US-10-279-733-10	Sequence 10, Appli

ALIGNMENTS

RESULT 1

US-10-041-775-2
; Sequence 2, Application US/10041775
; Publication No. US20030108564A1
; GENERAL INFORMATION:
; APPLICANT: BROWN, Eric L.
; APPLICANT: LEE, Lawrence
; APPLICANT: HOOK, Magnus
; TITLE OF INVENTION: METHOD OF PREVENTING T CELL-MEDIATED RESPONSES BY THE USE OF TH+
; TITLE OF INVENTION: HISTOCOMPATIBILITY COMPLEX CLASS II ANALOG PROTEIN (MAP PROTEIN)
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS
; FILE REFERENCE: P07023US01/BAS
; CURRENT APPLICATION NUMBER: US/10/041,775
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: 60/260,523
; PRIOR FILING DATE: 2001-01-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-041-775-2

Query Match 100.0%; Score 1016; DB 14; Length 200;
Best Local Similarity 100.0%; Pred. No. 1.7e-71;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MRGSHHHHGSQIPYITVNGTSONILSSLTFNKNQOISYKDIENKVKSVLYFNRGISD	60
DB	1	MRGSHHHHGSQIPYITVNGTSONILSSLTFNKNQOISYKDIENKVKSVLYFNRGISD	60
QY	61	IDRLRSKQAKTVHFKNGTKRWVDLKAGIHTADLINTSDIKAKISVNVDTKKQVKDKAKA	120
DB	61	IDRLRSKQAKTVHFKNGTKRWVDLKAGIHTADLINTSDIKAKISVNVDTKKQVKDKAKA	120

QY 121 NVQVPYITVNGTSQNILSLTFKKNQOISYKDLNNVSVLKNRGIITDVLRLSKQAK 180
Db 121 NVQVPYITVNGTSQNILSLTFKKNQOISYKDLNNVSVLKNRGIITDVLRLSKQAK 180
QY 181 FTVNFKNGTKKVIDLKAGIY 200
Db 181 FTVNFKNGTKKVIDLKAGIY 200

RESULT 2

US-09-982-992A-2
; Sequence 2, Application US/09982992A
; Patent No. US20020164337A1
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M. et al.
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE MAP PROTEIN AND METHOD OF USE IN TRE
; FILE REFERENCE: P06922US02/BAS
; CURRENT APPLICATION NUMBER: US/09/982,992A
; CURRENT FILING DATE: 2001-10-22
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/277,287
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/241,832
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-982-992A-2

Query Match 61.6%; Score 626; DB 9; Length 131;
Best Local Similarity 100.0%; Pred. No. 2.3e-41;
Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRGSHHHHGSQIPYITVNGTSQNILSLTFKKNQOISYKDLNNVSVLKNRGIISD 60
Db 1 MRGSHHHHGSQIPYITVNGTSQNILSLTFKKNQOISYKDLNNVSVLKNRGIISD 60
QY 61 IDRLSKQAKYTVHFKNGTKRVVDLKAGIHTADLINTSDIKAISVNVDTKKQVKDKEAKA 120
Db 61 IDRLSKQAKYTVHFKNGTKRVVDLKAGIHTADLINTSDIKAISVNVDTKKQVKDKEAKA 120

QY 121 NV 122
Db 121 NV 122

RESULT 3

US-10-041-775-4
; Sequence 4, Application US/10041775
; Publication No. US20030108564A1
; GENERAL INFORMATION:
; APPLICANT: BROWN, Eric L.
; APPLICANT: LEE, Lawrence
; APPLICANT: HOOK, Magnus
; TITLE OF INVENTION: METHOD OF PREVENTING T CELL-MEDIATED RESPONSES BY THE USE OF THE
; TITLE OF INVENTION: HISTOCOMPATIBILITY COMPLEX CLASS II ANALOG PROTEIN (MAP PROTEIN)
; FILE REFERENCE: P07023US01/BAS
; CURRENT APPLICATION NUMBER: US/10/041,775
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: 60/260,523
; PRIOR FILING DATE: 2001-01-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-041-775-4

Query Match 61.6%; Score 626; DB 14; Length 131;
Best Local Similarity 100.0%; Pred. No. 2.3e-41;
Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRGSHHHHGSQIPYITVNGTSQNILSLTFKKNQOISYKDLNNVSVLKNRGIISD 60
Db 1 MRGSHHHHGSQIPYITVNGTSQNILSLTFKKNQOISYKDLNNVSVLKNRGIISD 60
QY 61 IDRLSKQAKYTVHFKNGTKRVVDLKAGIHTADLINTSDIKAISVNVDTKKQVKDKEAKA 120
Db 61 IDRLSKQAKYTVHFKNGTKRVVDLKAGIHTADLINTSDIKAISVNVDTKKQVKDKEAKA 120
QY 121 NV 122
Db 121 NV 122

RESULT 4

US-08-781-986A-5238
; Sequence 5238, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 5238:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 150 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-781-986A-5238

Query Match 43.3%; Score 440; DB 8; Length 150;
Best Local Similarity 78.9%; Pred. No. 8.5e-27;
Matches 86; Conservative 13; Mismatches 10; Indels 0; Gaps 0;

QY 4 SHHHHHGSQIPYITVNGTSQNILSLTFKKNQOISYKDLNNVSVLKNRGIISD 63
Db 42 SLHHGSKVHVPAITVNGTSQNILSLTFKKNQOISYKDLNNVSVLKNRGIISD 101
QY 64 RLSKQAKYTVHFKNGTKRVVDLKAGIHTADLINTSDIKAISVNVDTKKQVKDKEAKA 112
Db 102 RLSKQAKYTVHFKNGTKRVVDLKAGIHTADLINTSDIKAISVNVDTKKQVKDKEAKA 150

```
RESULT 5
US-10-329-624-5238
; Sequence 5238, Application US/10329624
; Publication No. US20040043037A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
;   Gil H. Choi
;   Patrick S. Dillon
;   Craig A. Rosen
;   Steven C. Barash
;   Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 5233:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 151 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-781-986A-5233
; Query Match 20.7%; Score 210; DB 8; Length 151;
; Best Local Similarity 44.3%; Pred. No. 7.1e-09;
; Matches 43; Conservative 22; Mismatches 32; Indels 0; Gaps 0;
QY 14 IPYTTVNGTSQNLSSLTFFNKNQOISYKDIENKVKSVLYFNRGISDIDLRLSKQAKYTV 73
Db 55 VPTYIADVGINAFQNSYLNLPKDSQLSYLDLGNKVKALLYDERGYTPEKIRNAKSNVYTI 114
QY 74 HFKNGTKRVVDLKAGIHTADLINTSDIKAISVNVDTK 110
Db 115 TWKDGSKKEVDLKKDSYTNLFDSNSIKQIDINVTK 151
RESULT 6
US-10-329-624-5238
; Sequence 5238, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/329,624
; FILING DATE: 27-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/956,171
; FILING DATE: October 20, 1997
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248PDI1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 5238:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 150 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5238:
; Query Match 43.3%; Score 440; DB 15; Length 150;
; Best Local Similarity 78.9%; Pred. No. 8.5e-27;
; Matches 86; Conservative 13; Mismatches 10; Indels 0; Gaps 0;
QY 4 SHHHHHSQIPYTTVNGTSQNLSSLTFFNKNQOISYKDIENKVKSVLYFNRGISDIDL 63
Db 42 SLRHGYSKVHPVYAITVNGTSQNLSSLTFFNKNQOISYKDIENKVKSVLYFNRGISDIDL 101
QY 64 RLSQAKYTVHFNKGTKEVDLKAGIHTADLINTSDIKAISVNVDTKKQ 112
Db 102 RLSQAKYTVVFPNGTKKVIDLKAGIYTDLINTSEIKAININVDTKKQ 150
RESULT 7
US-10-329-624-5233
; Sequence 5233, Application US/10329624
; Publication No. US20040043037A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
;   Gil H. Choi
;   Patrick S. Dillon
;   Craig A. Rosen
;   Steven C. Barash
;   Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/329,624
; FILING DATE: 27-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/956,171
; FILING DATE: October 20, 1997
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 5233:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 151 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5233:
US-10-329-624-5233

Query Match 20.7%; Score 210; DB 15; Length 151;
Best Local Similarity 44.3%; Pred. No. 7.1e-09;
Matches 43; Conservative 22; Mismatches 32; Indels 0; Gaps 0;

QY 14 IPYITVNGTSONILSSITFNKQOISYKDIENKVKSVLYFNRGISDIDLRLSKOAKYTV 73
DB 55 VPYITVDGIMAFNOSYLNLPKDSQSLYDLGNKVKALLYDERGVTPKIRNAKSAVYTI 114

QY 74 HFKNGTKRVVDLKGAGHTADLINTSDIKAISVNVDTK 110
DB 115 TWKDSKKEVDLKKDSYANLPFDSNSIKQIDINVTK 151

RESULT 8
US-10-470-048B-221
; Sequence 221, Application US/10470048B
; Publication No. US20050037444A1
; GENERAL INFORMATION:
; APPLICANT: MEINKE ET AL.
; TITLE OF INVENTION: A METHOD FOR IDENTIFICATION, ISOLATION AND PRODUCTION OF
; FILE REFERENCE: SONN:035US
; CURRENT APPLICATION NUMBER: US/10/470,048B
; CURRENT FILING DATE: 2003-07-25
; NUMBER OF SEQ ID NOS: 603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 221
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-470-048B-221

Query Match 14.0%; Score 142; DB 17; Length 141;
Best Local Similarity 33.0%; Pred. No. 0.0013;
Matches 31; Conservative 25; Mismatches 38; Indels 0; Gaps 0;

QY 14 IPYITVNGTSONILSSITFNKQOISYKDIENKVKSVLYFNRGISDIDLRLSKOAKYTV 73
DB 47 VPYITVDGIMAFNOSYLNLPKDSQSLYDLGNKVKALLYDERGVTPKIRNAKSAVYTI 106

QY 74 HFKNGTKRVVDLKGAGHTADLINTSDIKAISVNV 107
DB 107 TLNDGNKVKVNLKNDDAKNSIDPSTIKQIQIVV 140

RESULT 9
```

```

US-10-378-674-4
; Sequence 4, Application US/10378674
; Publication No. US20040006209A1
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M.
; TITLE OF INVENTION: MONOCLONAL AND POLYCLONAL ANTIBODIES RECOGNIZING COAGULASE-NEGATI
; FILE REFERENCE: P07556US01/BAS
; CURRENT APPLICATION NUMBER: US/10/378,674
; CURRENT FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: 60/361,324
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-378-674-4

Query Match 13.1%; Score 133; DB 15; Length 343;
Best Local Similarity 22.4%; Pred. No. 0.02;
Matches 54; Conservative 32; Mismatches 71; Indels 84; Gaps 9;

QY 1 MRGSHHHHHGSGQIP-----YITVNGTS-----QNILSSLTF 33
DB 1 MRGSHHHHHGSLVPRGSEQGSNVNHLIKVTQDSITEGYDSDGIIKAHDAENLIYDVTF 60

QY 34 NKQOISYKDIENKVKSVLYFNRGISDIDL-----RLS 66
DB 61 -----EVDKVKSGDWTWNIDKNVTWPSDUTDSFAIPKIDNNGEIIATGYNTN 111

QY 67 KQAKYTVHFNGTKRVVDLKGAGHTADLINTSDIKAISVNVDTKKQVKDKREKANVQVPY 126
DB 112 KQIYTV--FTDYVDKYENIKAHKLKLTSDYIDSKVP-----NNNTKLDVEYKTALSSVNKTI 165

QY 127 TITVNGTSQNILSNLTF-----KQOQIS-----YKLENNKSVLKSNGRIGTDV 171
DB 166 TVEYQKPNENRTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVNISGNGEGSTII 225

QY 172 D 172
DB 226 D 226

RESULT 10
US-10-378-674-6
; Sequence 6, Application US/10378674
; Publication No. US20040006209A1
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M.
; TITLE OF INVENTION: MONOCLONAL AND POLYCLONAL ANTIBODIES RECOGNIZING COAGULASE-NEGATI
; FILE REFERENCE: P07556US01/BAS
; CURRENT APPLICATION NUMBER: US/10/378,674
; CURRENT FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: 60/361,324
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-378-674-6

Query Match 12.9%; Score 131; DB 15; Length 316;
Best Local Similarity 22.6%; Pred. No. 0.025;
Matches 53; Conservative 33; Mismatches 71; Indels 78; Gaps 9;

QY 1 MRGSHHHHHGSGQ-----IPYITVNGTS-----QNILSSLTFNKQOI 39
DB 1 MRGSHHHHHGSEQGSNVNHLIKVTQDSITEGYDSDGIIKAHDAENLIYDVTF----- 54
```



```
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 512 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-813-820-6

Query Match      11.5%; Score 116.5; DB 9; Length 512;
Best Local Similarity 23.8%; Pred. No. 0.63; Indels 97; Gaps 14;
Matches 64; Conservative 29; Mismatches 79;

QY 1 MRGSHHHHHGSGIPYITVNGTSQNLSS-----LTF-NKQQISYKD----- 43
Db 1 MRGSHHHHHGSGARDISST-NVTDLTVPSPKIEDGGKTVKMTFDDKNGKIQNGDMIKVA 59
QY 44 -----IENKVS-----LYFN-----RGISDIDRLSKQAKYT 72
Db 60 WPTSGTVKIEGYSKTVPLTVKGEVQAVITPDGATITFNDKVEKLSV-----SGFAEFE 115
QY 73 VHFPGTKRVVDLKGAGHTADLINTSDIKAISV---NVDTKQVXDKEA----- 118
Db 116 VQGRNLQ-----TNTSDKVATITSGNSTVTVHKSEAGTSSVFFYKGTG 161
QY 119 -----KANQVQVYITVNGTSQNLSTFKK-----NQGISYKDLNNV-----KSVLKSNR 166
Db 162 DMLPEDTTHRVFLNNEKSYSVDITIKQIQGGQQLDLSTLNINVTGTHSNYYSQS 221
QY 167 GITVDRLSKQAKFTVAFKNGTKKVIDL 195
Db 222 AITDFE-KAFPGSKITV---DNTKNTIDV 246

RESULT 14
US-10-282-122A-72043
; Sequence 72043, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Lianguu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Cart, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
```

```
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72043
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-282-122A-72043

Query Match      11.3%; Score 115; DB 15; Length 374;
Best Local Similarity 24.2%; Pred. No. 0.55; Indels 58; Gaps 13;
Matches 52; Conservative 42; Mismatches 63;

QY 18 ITVNGTSQNLSSITFNKNQOISYKDIENKVSIVY----FNRGISDIDRLSKQAKYTV 73
Db 131 ITVSGNKNAIESQLI----BELGIKK-SDYLTTLFQANRPERNLKSD-KWVKEALVY 184
QY 74 HFKNG-TKRVVD-----LKAGHTADLINTSDIKA--ISVNVTKKQVK 114
Db 185 HFPNHFTLRVKEYRIIAYRQTDKGYVPILENGTR-VDTVNASELPGSFVTINLDOEKEVR 243
QY 115 -----DKEAKANVQVPYITVNGTSQNLSTFNKNQOISYKDLNNKSVLKS 165
Db 244 ELVOKLAKLKLVLGSRKV--ISSVNSSTYKDLLLEKNK-----NSRVPL--- 289
QY 166 RGITVDRLSKQAKFTVFNKNGTKKVIDLKAGYI 200
Db 290 ---SEIDTKLPYYSKIKKNTLDGS--IVDMEVGIV 319

RESULT 15
US-10-057-531A-3
; Sequence 3, Application US/10057531A
; Publication No. US20030161838A1
; GENERAL INFORMATION:
; APPLICANT: Lyon, Jeffrey A.
; APPLICANT: Angov, Evelina
; TITLE OF INVENTION: Isolation and Purification of P. falciparum Merozoite
; TITLE OF INVENTION: Protein-142 Vaccine
; FILE REFERENCE: 003/241/SAP
; CURRENT APPLICATION NUMBER: US/10/057,531A
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 60/264,535
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/347,564
; PRIOR FILING DATE: 2001-10-26
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 3
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: E. coli expressed P. falciparum MSP142 (3D7) Protein
; OTHER INFORMATION: Sequence in pET42A
US-10-057-531A-3

Query Match      11.3%; Score 114.5; DB 14; Length 393;
Best Local Similarity 22.5%; Pred. No. 0.64;
Matches 53; Conservative 34; Mismatches 90; Indels 59; Gaps 10;

QY 4 SHHHHHH-GSQIP--YITVNGTSQNLSSITFNKNQOISY-----KDIENKV 48
Db 2 AHHHHHPGSGIEGRTMAISVTMDNILSG--FENEYDVLYLKLPLAGVYSLKKQIE--- 56
QY 49 KSVLYFNRGISDIDRLSKQAKYTVHFKNGTKRVVDLKGAGHTADLINTSDIKAISVND 108
Db 57 KNITFTNLNLDILNSRLKRGKRYL-----DVLESOLMQFKHSSNEYIIE 102
```

```
QY 109 TKQVKDKKAKANQVPTITVNGTSQNT---LSNLTFFKKNQOISYKDLNNVKSVLKSN 165
Db 103 DSFKLLNSEQKNTLLKSYKIKESVENDIKFAQEGISYKVKYKDLKESIKKVIKEE 162
QY 166 R-----GITDVLRLSKQAFTVN---FKNGTKKV-----IDLKAGI 199
Db 163 KEKFPSSPPTTPPSAKTDEQKESKFLPFLNIETLYNNLVNKIDDDYLINLKAKI 218

RESULT 16
US-10-057-532A-3
; Sequence 3, Application US/10057532A
; Publication No. US20030161839A1
; GENERAL INFORMATION:
; APPLICANT: Lyon, Jeffrey A.
; APPLICANT: Angov, Evelina
; APPLICANT: Cohen, Joe D.
; APPLICANT: Voss, Gerald
; TITLE OF INVENTION: Recombinant P. falciparum Merozoite Protein-142 Vaccine
; FILE REFERENCE: 003/238/SAP
; CURRENT APPLICATION NUMBER: US/10/057,532A
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: US 60/264,535
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/347,564
; PRIOR FILING DATE: 2001-10-26
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 3
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: E. coli expressed P. falciparum MSP142 (3D7) Protein Sequence in
US-10-057-532A-3

Query Match 11.3%; Score 114.5; DB 14; Length 393;
Best Local Similarity 22.5%; Pred. No. 0.64;
Matches 53; Conservative 34; Mismatches 90; Indels 59; Gaps 10;

QY 4 SHHHHHH-GSQIP--YITVNGTSQNLSSLLTFNKNQOISY-----KDIENKV 48
Db 2 AHHHHHHFGSGTEGRTMAISVTMDNLISG--FENEYDVIYKPLAGVRSKKQIE--- 56
QY 49 KSVLNFNRGISDIDRLRLSKQAQYVTHFKNGTKRVVDLKAGIHTADLINTSDIKAISVND 108
Db 57 KNIFTNLNLDILNSRLKRRKRYFL-----DVLESDLMPKHISSNEYIIE 102
QY 109 TKQVKDKKAKANQVPTITVNGTSQNT---LSNLTFFKKNQOISYKDLNNVKSVLKSN 165
Db 103 DSFKLLNSEQKNTLLKSYKIKESVENDIKFAQEGISYKVKYKDLKESIKKVIKEE 162
QY 166 R-----GITDVLRLSKQAFTVN---FKNGTKKV-----IDLKAGI 199
Db 163 KEKFPSSPPTTPPSAKTDEQKESKFLPFLNIETLYNNLVNKIDDDYLINLKAKI 218

RESULT 17
US-09-813-820-4
; Sequence 4, Application US/09813820
; Patent No. US20020102262A1
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Patti, Joseph M.
; APPLICANT: House-Pompeo, Karen
; APPLICANT: Schanam, Narayana
; APPLICANT: Symerky, Jindrich
; TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
AND METHODS OF USE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
```

```
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/813,820
; FILING DATE: 22-Mar-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/856,253
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: TAWK:193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 211 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-813-820-4

Query Match 11.2%; Score 114; DB 9; Length 211;
Best Local Similarity 23.9%; Pred. No. 0.32;
Matches 54; Conservative 28; Mismatches 84; Indels 60; Gaps 10;

QY 1 MRGSHHHHHGSGQPYTITVNGTSQNLSSLLTFNKNQOISYKDIENKSVLYFNRGISD 60
Db 1 MRGSHHHHHGSGDDKVAITISG-----NKSTNTVTHKSEAGTSSVFYKGTG--- 46
QY 61 IDRLSKQAQYVTHFKNGTKRVVDLKAGIHTADLINTS---DIKAISVNV----- 107
Db 47 -DM-LPETTHRVWFLNINNEKSYVSKDITTKDQIQGGQQLDLSTLINVTGTHSNYSG 104
QY 108 -----DTKQVKDKKAKANQVPTITVNGTSQNLSSLLTFNKNQOISYKDIENKSVLYFNRGISD 148
Db 105 QSAITDFEKAFFGSKITVDNTKNTIDVTIPQGYGSYNSFSINYKTKITNEQOQKEFVNSQ 164
QY 149 ISYKD---LENNVKSV-----LKSNRGITDV---DLRLSKQAQFT 182
Db 165 AMYQEHGKEEVNGKSFNHTVHNINANAGIEGTGKGLKVLKQDKDT 210

RESULT 18
US-10-282-122A-53616
; Sequence 53616, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
```

```

; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53616
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Clostridium difficile
US-10-282-122A-53616

Query Match      11.2%; Score 113.5; DB 15; Length 246;
Best Local Similarity 22.0%; Pred. No. 0.43;
Matches 44; Conservative 40; Mismatches 65; Indels 51; Gaps 10;

QY 23 TSONLSSLTFNKNQOI---SYKDLENKYSVLYFNRGISDIDLRLSKOAKYTVHFKNGT 79
DB 49 TKSNTMKELNVNLNENIFAYNFKMKNK---LKNPVYENVEIKRPLNKKIISLKE-- 102
QY 80 KKV-----VDLKAGHTADLINTSIKA-----ISVNDTKQVKDKKAKANV 122
DB 103 KEIFAVLKDEDNYCYIDKKGNI--LEELRGSNESKDLIVDDYSIDDKNSIKFKNYK--- 158
QY 123 QVPYTTVNGTSONLSNLTFKNQO---QISYKDL--ENNVSVLKSNRGITDVLRLS 176
DB 159 -----TKENVFKTLNLYKEGIYKYNVNLKESNIEMLTRSN-----IKILLS 203
QY 177 KOAKFTVNFKNGTKKVIDLK 196
DB 204 NDDNLDYNSRVSKILIDIQ 223

```

```

RESULT 19
US-10-282-122A-61097
; Sequence 61097, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELINRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078

```

```

; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61097
; LENGTH: 963
; TYPE: PRT
; ORGANISM: Legionella pneumophila
US-10-282-122A-61097

Query Match      10.9%; Score 111; DB 15; Length 963;
Best Local Similarity 21.6%; Pred. No. 3.7;
Matches 51; Conservative 46; Mismatches 73; Indels 66; Gaps 9;

QY 13 QIPYITVY-CTSONLSLTFNKNQOISYKDIENKVK----- 49
DB 119 KIETLLTVNKKITELIAETDLARQYQLLNDVNLKLVKXANFDLKEKLSIKAIRAKL 178
QY 50 ----SVLYFNRGISDID----LRLSKOAKYTVHFKNGTKRVDLKGAGHTAD----- 93
DB 179 NKEINILYEN-AVSSQDKKQPELTISEQIOYEANLYINNORIAVTHQKLNALQLKKIIK 237
QY 94 ----LINTSDIKAISVNDTKQVKDKREAKANVQPYTITVNGTSONLSNLTFKN--- 146
DB 238 TNLILLNLLDTRKALQSADINDYREVINOYS-----AIESLQTLNLSLTNQANLNV 287
QY 147 ---QOISYKDLNNVSVLKSNRGITDVLRLSKOAKFTVNFKNGTKKVIDLKAGI 199
DB 288 EPGLQKSVSLQNDVKTELKET---SNOKLILSKNLE---DYONQLKKQISTROSL 337

```

```

RESULT 20
US-10-193-764-28
; Sequence 28, Application US/10193764
; Publication No. US20030133943A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
; FILE REFERENCE: 1038-1239MIS
; CURRENT APPLICATION NUMBER: US/10/193,764
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: 09/167,568
; PRIOR FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 1220
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-193-764-28

Query Match      10.8%; Score 110; DB 14; Length 1220;

```

Best Local Similarity 23.7%; Pred. No. 6;
Matches 54; Conservative 28; Mismatches 82; Indels 64; Gaps 9;

Qy 3 GSHHHHGSQIPYTI-----TVNGTSQNLSSLTFFKQIISYKDIEKVKSV 51
Db 422 GSDFDNH---QKPLTIKDVINSGLTAGGVNINGLNLTVNGANL--KAITNFTFNV 476

Qy 52 --LYFNRGISDIDLRLSKQAKYTVHFKNKTRKRVVDLKAGIHTADLINTSDIKAISVNDT 109
Db 477 GGLFDNKGNSISI-----ARGGAKFKDINNTSSLN-ITNSTD 514

Qy 110 KKQ-----VKDKEAKANVQVPYTI-----TVNGTSQNLSSLTFFK-- 145
Db 515 TYRTIIEGNTWKAGDLNIIDNKGNAEIQIGNISQEGNLTISSDKINITQIIRKGV 574

Qy 146 NQOISYKDLNNVKSVLKSNRGITDVLRLSKQAKFTVNFKNGTKKVI 193
Db 575 NKEDSDSSANNANLTIKTELQLTGDLNISGFDKAEITAKEGADLII 622

RESULT 21
US-10-193-764-26
; Sequence 26, Application US/10193764
; Publication No. US20030133943A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
; FILE REFERENCE: 1038-1239MIS
; CURRENT APPLICATION NUMBER: US/10/193,764
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: 09/167,568
; PRIOR FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 1226
; TYPE: PRF
; ORGANISM: Haemophilus influenzae
US-10-193-764-26

Query Match 10.8%; Score 110; DB 14; Length 1226;
Best Local Similarity 23.7%; Pred. No. 6;
Matches 54; Conservative 28; Mismatches 82; Indels 64; Gaps 9;

Qy 3 GSHHHHGSQIPYTI-----TVNGTSQNLSSLTFFKQIISYKDIEKVKSV 51
Db 428 GSDFDNH---QKPLTIKDVINSGLTAGGVNINGLNLTVNGANL--KAITNFTFNV 482

Qy 52 --LYFNRGISDIDLRLSKQAKYTVHFKNKTRKRVVDLKAGIHTADLINTSDIKAISVNDT 109
Db 483 GGLFDNKGNSISI-----ARGGAKFKDINNTSSLN-ITNSTD 520

Qy 110 KKQ-----VKDKEAKANVQVPYTI-----TVNGTSQNLSSLTFFK-- 145
Db 521 TYRTIIEGNTWKAGDLNIIDNKGNAEIQIGNISQEGNLTISSDKINITQIIRKGV 580

Qy 146 NQOISYKDLNNVKSVLKSNRGITDVLRLSKQAKFTVNFKNGTKKVI 193
Db 581 NKEDSDSSANNANLTIKTELQLTGDLNISGFDKAEITAKEGADLII 628

RESULT 22
US-10-404-667-5
; Sequence 5, Application US/10404667
; Publication No. US20040005332A1
; GENERAL INFORMATION:
; APPLICANT: Angov, Evelina
; APPLICANT: Lyon, Jeffrey A.
; APPLICANT: Darko, Christian Asare
; APPLICANT: Cohen, Joe D.

; TITLE OF INVENTION: Recombinant P. falciparum Merozoite Protein-142 Vaccine
; FILE REFERENCE: 003/261/SAP
; CURRENT APPLICATION NUMBER: US/10/404,667
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/369,741
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/379,688
; PRIOR FILING DATE: 2002-05-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Microsoft Work XP
; SEQ ID NO 5
; LENGTH: 371
; TYPE: PRF
; ORGANISM: P. falciparum (FVO) MSP1-42
; FEATURE:
US-10-404-667-5

Query Match 10.7%; Score 108.5; DB 15; Length 371;
Best Local Similarity 21.7%; Pred. No. 1.8;
Matches 60; Conservative 39; Mismatches 79; Indels 99; Gaps 15;

Qy 4 SHHHHHH-GSQIPYTIITVNGTSQNLSSLTFFKQIISY-----KDIENKVK 50
Db 2 AHHHHHFGSGSGTAVTPSVIDNLSKI--ENEYEVLYKPLAGVYRSLKKQLN---N 56

Qy 51 VLYFNRGISDI-DLRLSKQ-----AKYTV-----HFKNKTRKRVVDLK 86
Db 57 VMTFNANVKDILNSRFNKNVLESDLIPYKDLTSSNVVVKDPYKFLNKEKDKFLS 116

Qy 87 AGIHTADLINT-----SDIKAISVNDTKQVQKDKKANVQVPY 126
Db 117 SYNYIKDSIDTIDINFANDVLGYKVKLSSEKYSKSDLSI-----KKYINDKQGENEKYLPF 170

Qy 127 TITVNGTSQNI-----LSNLTFFK-NQOISYKDLNNVKS-----LKS 165
Db 171 LNNIETLYKTVDKIDLFVHLEAEVLNLYTEKSNVEVKIKEL-NYLTQIDKLADFKCN 229

Qy 166 R---GITVDL-----RLSKQAKFTVNFKNGTKKVI 193
Db 230 NNFVGIADLTDYNNHNNLLTKFLSTGMVFENPAKTVL 266

RESULT 23
US-09-813-820-7
; Sequence 7, Application US/09813820
; Patent No. US20020102262A1
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; Patti, Joseph M.
; House-Pompeo, Karen
; Sthanam, Narayana
; Symersky, Jindrich
; TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESS: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/813,820
; FILING DATE: 22-Mar-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: 08/856,253


```
US-10-282-122A-70558
; Sequence 70558, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangau
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70558
; LENGTH: 1051
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-282-122A-70558

Query Match      10.2%; Score 104; DB 15; Length 1051;
Best Local Similarity 25.9%; Pred. No. 15;
Matches 53; Conservative 29; Mismatches 69; Indels 54; Gaps 11;

QY 25 QNLSLTFNKNQOISYKDIENKVKSVLYFNRGISDIDLRLSKQAKYTVHFKNGT----- 79
DB 50 QTTMGATPQSTQDSISSKIDNQVSLAYVNSVQTE---SIPNASIVTVEYDNGTDMKA 106
QY 80 ----KRVVD---LKAGIHTADLINTS-----DIKAIISVNVDTKKQVKDKKAKANVQ-VPYT 127
DB 107 EQQLKKEIDKIKFDGVEPELTRNSMDAPPVAVSFTSNNQK-LKDYTKLNLQQLVPKL 165
QY 128 ITVNTGTSQNLSNLTFTKNQOISYK-----DLNNVKSVL----- 162
DB 166 QTIDGV-QN--AQLNGQTNREVSFLFKQKNLDEKGLTANDVENYIKTATRETPPLGLFQFN 222
QY 163 KSNRGITVDLRLSKQAKFTVNFKN 187
DB 223 KSNKSIV-----VDGQFKSVDAFKN 242

RESULT 27
US-10-369-493-2021
; Sequence 2021, Application US/10369493

US-10-282-122A-60961
; Sequence 60961, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09

Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2021
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-2021

Query Match      10.2%; Score 103.5; DB 15; Length 609;
Best Local Similarity 24.9%; Pred. No. 8;
Matches 51; Conservative 29; Mismatches 56; Indels 69; Gaps 11;

QY 24 SQN--ILSSLTFNKNQOISYKDI--EN-KVKSVL-----YFNRGISDIDLRLSKQAKYT 72
DB 302 SQNDKILKLETTNKAYTKYKEVSLKKEAFKELDNESYNH-----DEBLKKYKYT 356
QY 73 VHFKNGTKRVDLKGAIHTADLINTSDIKATSVNVDTKKQVKDKKAKAN----- 121
DB 357 -----RETLDRVNRREQLIIDQNEFLKKSVMNQLQNEVNTATNFKSFLKE 400
QY 122 --VQVPTTYTNGTS-----QNLSNLTFTKNQO-----ISYKOLENNKSVLKSNGRITDV 171
DB 401 KYAKLADSITELNTSTKKEALGENTLFECHLEICLKYK---KNENISNTNKLQN- 456
QY 172 DLRLSKQAKFTVNFKNQTKKVIDLK 196
DB 457 -----SFKNERKKVLDLR 469

RESULT 28
US-10-282-122A-60961
; Sequence 60961, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
```

```
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60961
; LENGTH: 1778
; TYPE: PR1
; ORGANISM: Listeria monocytogenes
US-10-282-122A-60961

Query Match      10.2%; Score 103.5; DB 15; Length 1778;
Best Local Similarity 20.2%; Pred. No. 31;
Matches 50; Conservative 38; Mismatches 73; Indels 87; Gaps 11;

QY      17 TTTVNGTSONILSLTFKNNQOISYKDIENKVKSVLYENRSGISDID-----62
Db      1014 TTVNKDPAPISAKT-----EITDKSKKTEAAF-----LDDIDADTNDGSI VTSNFA 1063

QY      63 --LRLSKQAKYTVHPKNGTKRVVDLKAGIHTADLINTSDIKAISVNVDTKKQ-----112
Db      1064 TAVNLDKAGDYTVTLNSINS---DGVAGTPTAIIVHVEKEKIATISTNTAQOYKYAKIN 1120

QY      113 ----VKKEKAN-----VQVP--YITVNGTSON-----136
Db      1121 ETQFLKDVHASINASPPTAVLESDETGVKLDVPGTYTITATNEDGGVSAPEVSVIV 1180

QY      137 -----ILSLTFKNNQOISYKDLNNKSVLK--SNRGIT----DVDLRLSKQAKFTV 183
Db      1181 RKIPAPEITADKEIITPKFDEVSEAEFLNDIHATISDKNVAITSNFTSDVNLNKAGDYTV 1240

QY      184 NPKNGTKK 191
Db      1241 TL-NATNE 1247

RESULT 29
US-09-345-373-30
; Sequence 30, Application US/09345373
; Publication No. US20030077695A1
; GENERAL INFORMATION:
; APPLICANT: RUBEN, STEVEN M.
; APPLICANT: JIMENEZ, PABLO
; APPLICANT: DUAN, D. ROXANNE
; APPLICANT: RAMPY, MARK A.
; APPLICANT: MENDRICK, DONNA
; APPLICANT: ZHANG, JUN
; APPLICANT: NI, JIAN
; APPLICANT: MOORE, PAUL A.
; APPLICANT: COLEMAN, TIMOTHY A.
; APPLICANT: GRUBER, JOACHIM R.
; APPLICANT: DILLON, PATRICK J.
; APPLICANT: GENTZ, REINER L.
; TITLE OF INVENTION: KERATINOCYTE GROWTH FACTOR-2
; NUMBER OF SEQUENCES: 148
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 NEW YORK AVE, NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: /09/345,373
; APPLICATION NUMBER: US/09/345,373
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/023,082
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/461,195
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/023,852
; FILING DATE: 13-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/039,045
; FILING DATE: 28-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/862,432
; FILING DATE: 23-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/910,875
; FILING DATE: 13-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/055,561
; FILING DATE: 13-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFEE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0360008/EKS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 184 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-345-373-30

Query Match      10.1%; Score 102.5; DB 10; Length 184;
Best Local Similarity 24.3%; Pred. No. 2.1;
Matches 41; Conservative 26; Mismatches 61; Indels 41; Gaps 8;

QY      1 MRGSHHHHGSQIPYITVNGTSONILSLTFKNNQOISYKDIENKVKSVLYFN--RGI 58
Db      1 MRGSHHHHHS-----COALGQDMVSPPEATNSSSS--SFSPPSSAGRHVRSYNHLOG- 51

QY      59 SDIDLRLSKQAKYTVHF-----KN-----GTRK-----VVDLKAGIHTADLINTSDIK 101
Db      52 ---DVRWRKLFSTFKYFKLIEKNKGVSGTKENCPYSILEITSVEIGVAVKAINSYYL 108

QY      102 AIS-----VNVDTKQVKDKKAKANQVQPYTITVNGTSONILSN 140
Db      109 AMNKKGLYSGKEFNNDCKLKERIEINGYNTVYASFNMQHNGRQMYVALN 157

RESULT 30
US-10-075-446-30
; Sequence 30, Application US/10075446
; Publication No. US20030129687A1
; GENERAL INFORMATION:
; APPLICANT: RUBEN, STEVEN M.
; APPLICANT: JIMENEZ, PABLO
; APPLICANT: DUAN, D. ROXANNE
; APPLICANT: RAMPY, MARK A.
; APPLICANT: MENDRICK, DONNA
; APPLICANT: ZHANG, JUN
; APPLICANT: NI, JIAN
; APPLICANT: MOORE, PAUL A.
; APPLICANT: COLEMAN, TIMOTHY A.
```



```
/ GRUBER, JOACHIM R.
/ TITLE OF INVENTION: KERATINOCYTE GROWTH FACTOR-2
/ NUMBER OF SEQUENCES: 148
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
/ STREET: 1100 NEW YORK AVE, NW, SUITE 600
/ CITY: WASHINGTON
/ STATE: DC
/ COUNTRY: USA
/ ZIP: 20005-3934
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/075,446
/ FILING DATE: 15-Feb-2002
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 09/023,082
/ FILING DATE: <Unknown>
/ APPLICATION NUMBER: PCT/US95/01790
/ FILING DATE: 14-FEB-1995
/ APPLICATION NUMBER: US 08/461,195
/ FILING DATE: 05-JUN-1995
/ APPLICATION NUMBER: US 60/023,852
/ FILING DATE: 13-AUG-1996
/ APPLICATION NUMBER: US 60/039,045
/ FILING DATE: 28-FEB-1997
/ APPLICATION NUMBER: US 08/862,432
/ FILING DATE: 23-MAY-1997
/ APPLICATION NUMBER: US 08/910,875
/ FILING DATE: 13-AUG-1997
/ APPLICATION NUMBER: US 60/055,561
/ FILING DATE: 13-AUG-1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: STEFFEE, ERIC K.
/ REGISTRATION NUMBER: 36,688
/ REFERENCE/DOCKET NUMBER: 1488.0360008/EKS
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-371-2600
/ TELEFAX: 202-371-2540
/ INFORMATION FOR SEQ ID NO: 30:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 184 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-10-075-446-30
```

```
Query Match          10.1%; Score 102.5; DB 14; Length 184;
Best Local Similarity 24.3%; Pred. No. 2.1;
Matches 41; Conservative 26; Mismatches 61; Indels 41; Gaps 8;

QY 1 MRGSHHHHGHGSIPIYTTIVNGTSQNILSSLTFFNKNQOISYKDIENKVKSVLYFN--RGI 58
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1 MRGSHHHHGHGSIPIYTTIVNGTSQNILSSLTFFNKNQOISYKDIENKVKSVLYFN--RGI 51

QY 59 SDIDLRLSKQAKYTVHF-----KN-----GTKR-----VVDLKAGIHTADLINTSDIK 101
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
52 ---DVRWRKLFSTFKYFLKIEKNGKVGSKTKENCPCYSILEITSVIGVAVKAINSNYYL 108

QY 102 AIS-----VNVDTKKQVKDEAKANVQVPYTTIVNGTSQNILSN 140
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
109 AMNKKGLYGSKEFNNDCKLKERIEENGYNITYASFVWQHNGRQMYVALN 157
```

```
RESULT 31
US-10-035-212-30
; Sequence 30, Application US/10035212
; Publication No. US200301869041
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Ruben, Steven M.
/ APPLICANT: Jimenez, Pablo
/ APPLICANT: Duan, D. Roxanne
/ APPLICANT: Rampy, Mark A.
/ APPLICANT: Mendrick, Donna
/ APPLICANT: Zhang, Jun
/ APPLICANT: Ni, Jian
/ APPLICANT: Moore, Paul A.
/ APPLICANT: Coleman, Timothy A.
/ APPLICANT: Gruber, Joachim R.
/ APPLICANT: Dillon, Patrick J.
/ APPLICANT: Gentz, Reiner L.
/ TITLE OF INVENTION: Keratinocyte Growth Factor-2
/ FILE REFERENCE: 1488.0360000
/ CURRENT APPLICATION NUMBER: US/10/035,212
/ CURRENT FILING DATE: 2002-01-04
/ PRIOR APPLICATION NUMBER: 60/259,853
/ PRIOR FILING DATE: 2001-01-05
/ PRIOR APPLICATION NUMBER: 60/286,368
/ PRIOR FILING DATE: 2001-04-26
/ PRIOR APPLICATION NUMBER: 60/331,168
/ PRIOR FILING DATE: 2001-11-09
/ NUMBER OF SEQ ID NOS: 176
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 30
/ LENGTH: 184
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: pQE60-Cys37
/ OTHER INFORMATION: construct
US-10-035-212-30

Query Match          10.1%; Score 102.5; DB 14; Length 184;
Best Local Similarity 24.3%; Pred. No. 2.1;
Matches 41; Conservative 26; Mismatches 61; Indels 41; Gaps 8;

QY 1 MRGSHHHHGHGSIPIYTTIVNGTSQNILSSLTFFNKNQOISYKDIENKVKSVLYFN--RGI 58
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1 MRGSHHHHGHGSIPIYTTIVNGTSQNILSSLTFFNKNQOISYKDIENKVKSVLYFN--RGI 51

QY 59 SDIDLRLSKQAKYTVHF-----KN-----GTKR-----VVDLKAGIHTADLINTSDIK 101
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
52 ---DVRWRKLFSTFKYFLKIEKNGKVGSKTKENCPCYSILEITSVIGVAVKAINSNYYL 108

QY 102 AIS-----VNVDTKKQVKDEAKANVQVPYTTIVNGTSQNILSN 140
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
109 AMNKKGLYGSKEFNNDCKLKERIEENGYNITYASFVWQHNGRQMYVALN 157
```

```
RESULT 32
US-10-901-210-30
; Sequence 30, Application US/10901210
; Publication No. US20050037966A1
/ GENERAL INFORMATION:
/ APPLICANT: Ruben et al.
/ TITLE OF INVENTION: Keratinocyte Growth Factor-2
/ FILE REFERENCE: PFI55P2D1
/ CURRENT APPLICATION NUMBER: US/10/901,210
/ CURRENT FILING DATE: 2004-07-29
/ PRIOR APPLICATION NUMBER: 10/035,212
/ PRIOR FILING DATE: 2002-01-04
/ PRIOR APPLICATION NUMBER: 60/259,853
/ PRIOR FILING DATE: 2001-01-08
/ PRIOR APPLICATION NUMBER: 60/286,368
/ PRIOR FILING DATE: 2001-04-26
/ PRIOR APPLICATION NUMBER: 60/331,168
/ PRIOR FILING DATE: 2001-11-09
/ NUMBER OF SEQ ID NOS: 176
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 30
/ LENGTH: 184
```

```

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: p0860-Cys37 construct
US-10-901-210-30

Query Match      10.1%; Score 102.5; DB 17; Length 184;
Best Local Similarity 24.3%; Pred. No. 2.1; Indels 41; Gaps 8;
Matches 41; Conservative 26; Mismatches 61;

QY   1 MRGSHHHHGGSGIPVITVTNGTSQNILSSLTFFNNKQIQSYKDIEKNVKSVLYFN--RGI 58
     ||||| ||||| :||: ||||| :||: ||||| :||: ||||| :||: ||||| :||: |||||
DB   1 MRGSHHHHGGSS-----CQALGDQMSPEATNSSSS-SFSPSSAGRHVSRYNHLOG- 51

QY   59 SDIDLRLSQAOKYTVHP-----KN---GTKR-----VVLDKAGIHTADLINTSDIK 101
     ||||| ||||| :||: ||||| :||: ||||| :||: ||||| :||: ||||| :||: |||||
DB   52 ---DVRWRKLFSFTKYFLKIERNKGVSGTKENCPYSILEITSVEIGVVAVKAINSYYL 108

QY   102 AIS-----VNVDTKQVNDKEAKANVOVPYITVTNGTSQNILSN 140
     ||||| ||||| :||: ||||| :||: ||||| :||: ||||| :||: ||||| :||: |||||
DB   109 AMNKKGLYGSKFENNDCLEKRIENGNYTASFWHQHGRQMYYVALN 157

RESULT 33
US-10-404-667-2
; Sequence 2, Application US/10404667
; Publication No. US20040005332A1
; GENERAL INFORMATION:
; APPLICANT: Angov, Evelina
; APPLICANT: Lyon, Jeffrey A.
; APPLICANT: Darko, Christian Asare
; APPLICANT: Cohen, Joe D.
; TITLE OF INVENTION: Recombinant P. falciparum Merozoite Protein-142 Vaccine
; FILE REFERENCE: 003/261/SAP
; CURRENT APPLICATION NUMBER: US/10/404,667
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/369,741
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/379,688
; PRIOR FILING DATE: 2002-05-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Microsoft Work XP
; SEQ ID NO 2
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: P. falciparum (FVO) MSP1-42 (Mut 7, Pause Site mutant)
US-10-404-667-2

Query Match      10.0%; Score 101.5; DB 15; Length 371;
Best Local Similarity 22.7%; Pred. No. 6.2; Indels 101; Gaps 16;
Matches 63; Conservative 34; Mismatches 80;

QY   4 SHHHHH--GSGIPVITVTNGTSQNILSSLTFFNNKQIQSY-----XDIENKVX 49
     ||||| ||||| :||: ||||| :||: ||||| :||: ||||| :||: ||||| :||: |||||
DB   2 AHHHHHPPGGSGGTMAVTPSIDNILSKI--ENEYEVLKPLAGVYRSLLKOLEN--- 56

QY   50 SVLYFNRGISDT-DLRLSQ-----AKYTV----HFNGTKRKRVDDL 85
     ||||| ||||| :||: ||||| :||: ||||| :||: ||||| :||: ||||| :||: |||||
DB   57 NVMTFNVNVKIDILSRFRNFKNVLESDLIPYKDLTSSNVYVKDPYKFLNKERDKEL 116

QY   86 KAGIHTADLINF-----SDIKALSINVDTKKQVQKEAKA----- 120
     ||||| ||||| :||: ||||| :||: ||||| :||: ||||| :||: ||||| :||: |||||
DB   117 SSVNYIKDSIDFDINFANDVLGYKVLSEKYSKSLDSI-----KKYINDKQENKEYLPF 170

QY   121 --NVQPYTITVG-----TSQNILSNLTFFKNNQIQSYKDLENNVKSV-----LKS 164
     ||||| ||||| :||: ||||| :||: ||||| :||: ||||| :||: ||||| :||: |||||
DB   171 LNNIETLYK-TVNDKIDLFIHVHLEAEVLNTTYEKSNEVEVKIKEL-NYLKTIQDKLAFKK 228

QY   165 NR---GITDVLD-----RLSKQAOKFTVFNGTKTKVI 193
     ||||| ||||| :||: ||||| :||: ||||| :||: ||||| :||: ||||| :||: |||||

```



```
Qy 52 --LYFNRGISDI-----DLRLSKQAKYTVHFKNKGTKRVDLKGAIHTADLIN-TS 98
Db 918 GGLFDNKGNSISIAKGGARFKDIDNSKNLSITNSSSTYRTII-----SGNITNKG 970
Qy 99 DIKAISSVNDTKQ-----VKDKEAKANVQVPYITVNGTSQNLISLTFFKK--NQOISYK 152
Db 971 DLNITNEGSDTEMQIGGVSQKEG-----NLTISDKINITYKQITIKAGVDGENSDS 1022
Qy 153 DLNNVKSVLKSNRGITDVLRLSKQAKFTVNFKNKGTKKVIDLKAG 198
Db 1023 DATNNANLTIKTKEKLKLTQDLNLSGFKAEITAKDGS----DLTIG 1064
```

```
RESULT 39
US-10-193-764-63
; Sequence 63, Application US/10193764
; Publication No. US20030133943A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
; TITLE OF INVENTION: MOLECULAR WEIGHT PROTEINS
; FILE REFERENCE: 1038-1239MIS
; CURRENT APPLICATION NUMBER: US/10/193,764
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: 09/167,568
; PRIOR FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 1536
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-193-764-63
```

```
Query Match 9.9%; Score 101; DB 14; Length 1536;
Best Local Similarity 23.9%; Pred. No. 40;
Matches 54; Conservative 32; Mismatches 86; Indels 54; Gaps 11;

Qy 3 GSHHHHHHGSQIPYTI-----TVNGTSQNLISLTFFKNKQOISYKDIEKNKSV 51
Db 863 GSDFDNH---OKPLTIKKDVIINSGLTAGNIVNIAGNLTVESN--ANFKAITNFTFNV 917
Qy 52 --LYFNRGISDI-----DLRLSKQAKYTVHFKNKGTKRVDLKGAIHTADLIN-TS 98
Db 918 GGLFDNKGNSISIAKGGARFKDIDNSKNLSITNSSSTYRTII-----SGNITNKG 970
Qy 99 DIKAISSVNDTKQ-----VKDKEAKANVQVPYITVNGTSQNLISLTFFKK--NQOISYK 152
Db 971 DLNITNEGSDTEMQIGGVSQKEG-----NLTISDKINITYKQITIKAGVDGENSDS 1022
Qy 153 DLNNVKSVLKSNRGITDVLRLSKQAKFTVNFKNKGTKKVIDLKAG 198
Db 1023 DATNNANLTIKTKEKLKLTQDLNLSGFKAEITAKDGS----DLTIG 1064
```

```
RESULT 40
US-10-681-171-2
; Sequence 2, Application US/10681171
; Publication No. US20050053618A1
; GENERAL INFORMATION:
; APPLICANT: Barenkamp, Stephen J.
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT SURFACE PROTEINS OF NON-TYPEABLE
; FILE REFERENCE: 5346
; CURRENT APPLICATION NUMBER: US/10/681,171
; CURRENT FILING DATE: 2003-10-09
; PRIOR APPLICATION NUMBER: 09/155,614
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 08/617,697
; PRIOR FILING DATE: 1996-04-01
```

```
; PRIOR APPLICATION NUMBER: PCT/US97/04707
; PRIOR FILING DATE: 1997-04-01
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1536
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-681-171-2
```

```
Query Match 9.9%; Score 101; DB 17; Length 1536;
Best Local Similarity 23.9%; Pred. No. 40;
Matches 54; Conservative 32; Mismatches 86; Indels 54; Gaps 11;

Qy 3 GSHHHHHHGSQIPYTI-----TVNGTSQNLISLTFFKNKQOISYKDIEKNKSV 51
Db 863 GSDFDNH---OKPLTIKKDVIINSGLTAGNIVNIAGNLTVESN--ANFKAITNFTFNV 917
Qy 52 --LYFNRGISDI-----DLRLSKQAKYTVHFKNKGTKRVDLKGAIHTADLIN-TS 98
Db 918 GGLFDNKGNSISIAKGGARFKDIDNSKNLSITNSSSTYRTII-----SGNITNKG 970
Qy 99 DIKAISSVNDTKQ-----VKDKEAKANVQVPYITVNGTSQNLISLTFFKK--NQOISYK 152
Db 971 DLNITNEGSDTEMQIGGVSQKEG-----NLTISDKINITYKQITIKAGVDGENSDS 1022
Qy 153 DLNNVKSVLKSNRGITDVLRLSKQAKFTVNFKNKGTKKVIDLKAG 198
Db 1023 DATNNANLTIKTKEKLKLTQDLNLSGFKAEITAKDGS----DLTIG 1064
```

Search completed: March 16, 2005, 10:55:25
Job time : 139 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 16, 2005, 10:36:37 ; Search time 43 Seconds
(without alignments)
347.205 Million cell updates/sec

Title: US-10-041-775-2
Perfect score: 1016
Sequence: 1 MRGSHHHHGSQIPYITTV.....FTVFNKGTKKVIDLKAGIY 200

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS-COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	950	93.5	689	1	US-08-248-021A-2
2	440	43.3	150	4	US-08-956-171E-5238
3	440	43.3	150	4	US-08-781-986A-5238
4	210	20.7	151	4	US-08-956-171E-5233
5	210	20.7	151	4	US-08-781-986A-5233
6	159	15.6	31	1	US-08-248-021A-3
7	155	15.3	31	1	US-08-248-021A-4
8	116.5	11.5	512	3	US-08-856-253-6
9	114	11.2	211	3	US-08-856-253-4
10	110	10.8	31	1	US-08-248-021A-7
11	110	10.8	1095	4	US-09-206-942-45
12	110	10.8	1101	4	US-09-206-942-43
13	110	10.8	1220	4	US-09-206-942-28
14	110	10.8	1226	4	US-09-206-942-26
15	108.5	10.7	288	3	US-09-134-001C-4005
16	106	10.4	345	3	US-08-856-253-7
17	104	10.2	74	4	US-09-490-291-11
18	104	10.2	74	4	US-08-856-253-2
19	104	10.2	1051	3	US-09-134-001C-5005
20	103.5	10.2	609	4	US-09-538-092-711
21	102.5	10.1	184	3	US-09-023-082A-30
22	102.5	10.1	184	4	US-09-248-998-30
23	102.5	10.1	184	4	US-09-610-651-30
24	101.5	10.0	3241	4	US-09-841-786-1
25	101	9.9	1095	4	US-09-206-942-69
26	101	9.9	1536	1	US-08-038-682-2
27	101	9.9	1536	1	US-08-302-832-2

28	101	9.9	1536	2	US-08-530-198-2	Sequence 2, Appli
29	101	9.9	1536	2	US-08-469-880-2	Sequence 2, Appli
30	101	9.9	1536	2	US-08-728-470-2	Sequence 2, Appli
31	101	9.9	1536	2	US-08-617-697-2	Sequence 2, Appli
32	101	9.9	1536	3	US-08-719-641-2	Sequence 2, Appli
33	101	9.9	1536	4	US-09-206-942-67	Sequence 67, Appli
34	99.5	9.8	1207	4	US-09-978-594-4	Sequence 4, Appli
35	97.5	9.6	297	4	US-10-030-031A-2	Sequence 2, Appli
36	97.5	9.6	333	4	US-09-248-796A-15901	Sequence 15901, A
37	96.5	9.5	789	4	US-09-604-958-1	Sequence 1, Appli
38	96.5	9.5	789	4	US-09-995-587A-1	Sequence 1, Appli
39	95.5	9.4	221	1	US-08-621-081A-19	Sequence 19, Appli
40	95	9.4	331	4	US-09-107-433-3308	Sequence 3308, Ap
41	95	9.4	447	3	US-08-961-083-182	Sequence 182, App
42	95	9.4	447	3	US-09-536-784-182	Sequence 182, App
43	95	9.4	484	4	US-09-468-656A-6	Sequence 6, Appli
44	95	9.4	1039	4	US-09-583-110-5226	Sequence 5226, Ap
45	94.5	9.3	1231	4	US-09-071-035-420	Sequence 420, App

ALIGNMENTS

RESULT 1
US-08-248-021A-2
; Sequence 2, Application US/08248021A
; Patent No. 5648240
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Jonsson, Klas
; APPLICANT: Patti, Joseph M.
; APPLICANT: Gurusiddappa, Sivashankarappa
; TITLE OF INVENTION: MHC II ANALOG FROM STAPHYLOCOCCUS AUREUS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/248,021A
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: TANK:155
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 689 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-248-021A-2

Query Match 93.5%; Score 950; DB 1; Length 689;
Best Local Similarity 97.0%; Pred. No. 7.5e-82;
Matches 191; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 4 SHHHHGSQIPYITTVNGTSNTLSSTFNKNQOISYKDIENKVKSVLYFNRGISDIDL 63
DB 41 SLHHGYSKIQIPYITTVNGTSNTLSSTFNKNQOISYKDIENKVKSVLYFNRGISDIDL 100

QY 64 RLSKQAKYTVHFKNQTKRVVDLKGAGIHTADLINTSDIKAISVNVDTTKQVKDKKQKAKANVQ 123
DB 101 RLSKQAKYTVHFKNQTKRVVDLKGAGIHTADLINTSDIKAISVNVDTTKQVKDKKQKAKANVQ 160
QY 124 VPYITVNGTSONILSNLTFKKNQOISYKDLNNVKSVLKSNRGIITVDLRLSKQAKFTV 183
DB 161 VPYITVNGTSONILSNLTFKKNQOISYKDLNNVKSVLKSNRGIITVDLRLSKQAKFTV 220
QY 184 NFKNGTKKVIDLKAGIY 200
DB 221 NFKNGTKKVIDLKAGIY 237
RESULT 2
US-08-956-171E-5238
; Sequence 5238, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 5238:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 150 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5238:
US-08-956-171E-5238
Query Match 43.3%; Score 440; DB 4; Length 150;
Best Local Similarity 78.9%; Pred. No. 2.6e-34;
Matches 86; Conservative 13; Mismatches 10; Indels 0; Gaps 0;
QY 4 SHHHHHGSGIPYITVNGTSONILSNLTFKKNQOISYKDIENKVKSVLYFNRGISDIDL 63
DB 42 SLHHGYSKVHPVYAITVNGTSONILSNLTFKKNQOISYKDLDRVKSVLKSDRGISDIDL 101
QY 64 RLSKQAKYTVHFKNQTKRVVDLKGAGIHTADLINTSDIKAISVNVDTTKQ 112
DB 102 RLSKQAKYTVHFKNQTKRVVDLKGAGIHTADLINTSDIKAISVNVDTTKQ 150
RESULT 4
US-08-956-171E-5233
; Sequence 5233, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256

DB 102 RLSKQAKYTVHFKNQTKRVVDLKGAGIHTADLINTSDIKAISVNVDTTKQ 150
RESULT 3
US-08-781-986A-5238
; Sequence 5238, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 5238:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 150 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-781-986A-5238
Query Match 43.3%; Score 440; DB 4; Length 150;
Best Local Similarity 78.9%; Pred. No. 2.6e-34;
Matches 86; Conservative 13; Mismatches 10; Indels 0; Gaps 0;
QY 4 SHHHHHGSGIPYITVNGTSONILSNLTFKKNQOISYKDIENKVKSVLYFNRGISDIDL 63
DB 42 SLHHGYSKVHPVYAITVNGTSONILSNLTFKKNQOISYKDLDRVKSVLKSDRGISDIDL 101
QY 64 RLSKQAKYTVHFKNQTKRVVDLKGAGIHTADLINTSDIKAISVNVDTTKQ 112
DB 102 RLSKQAKYTVHFKNQTKRVVDLKGAGIHTADLINTSDIKAISVNVDTTKQ 150
RESULT 4
US-08-956-171E-5233
; Sequence 5233, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256


```

;
; TOPOLOGY: linear
; US-08-248-021A-3
;
; Query Match 15.6%; Score 159; DB 1; Length 31;
; Best Local Similarity 100.0%; Pred. No. 1.4e-08;
; Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 60 DIDLRLSKQAKYTVHFKNGTKRVVDLKGAGIH 90
; DB 1 DIDLRLSKQAKYTVHFKNGTKRVVDLKGAGIH 31
;
; RESULT 7
; US-08-248-021A-4
; Sequence 4, Application US/08248021A
; Patent No. 5648240
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Jonsson, Klas
; APPLICANT: Patti, Joseph M.
; APPLICANT: Gurusiddappa, Sivashankarappa
; TITLE OF INVENTION: MHC II ANALOG FROM STAPHYLOCOCCUS AUREUS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/248,021A
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: TAMK:155
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-248-021A-4
;
; Query Match 15.3%; Score 155; DB 1; Length 31;
; Best Local Similarity 100.0%; Pred. No. 3.5e-08;
; Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 170 DVDLRLSKQAKFTVAFKNGTKKVIDLKGAGIY 200
; DB 1 DVDLRLSKQAKFTVAFKNGTKKVIDLKGAGIY 31
;
; RESULT 8
; US-08-856-253-6
; Sequence 6, Application US/08856253
; Patent No. 6288214
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Patti, Joseph M.
; APPLICANT: House-Pompeo, Karen
; APPLICANT: Sthanam, Narayana
; APPLICANT: Symersky, Jindrich
; TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
;
; TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
; TITLE OF INVENTION: AND METHODS OF USE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/856,253
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/017,678
; FILING DATE: 16-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: TAMK:193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 512 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-856-253-6
;
; Query Match 11.5%; Score 116.5; DB 3; Length 512;
; Best Local Similarity 23.8%; Pred. No. 0.0085;
; Matches 64; Conservative 29; Mismatches 79; Indels 97; Gaps 14;
;
; QY 1 MRGSHHHHGGSQIPYITVNGTSQNILSS-----LTF-NKNQOISYKD----- 43
; DB 1 MRGSHHHHGGARDISST-NVTLTVSPSKIEDGGKTTVKTATPDNDKNGKIQNGDMIKVA 59
; QY 44 -----LENKVKSV-----LYEN---RGISDIDLRLSKQAKYT 72
; DB 60 WPTSGTVKIEGYSKTVPLTVKGEQVQAVTPDQATITFDNKVKELSDV-----SGFAEFE 115
; QY 73 VHFKNKGRVVDLKGAGIHTADLINTSDIKAISV---NVDTKKQVKQKEA----- 118
; DB 116 VQGRNLQ-----TNTSDKVATITSGNKSTNVTVHKSEAGTSSVFYKGTG 161
; QY 119 ----KANVQVPYITVNGTSQNILSNLTFKK---NQOI SYKOLENNV-----KSVLKSNR 166
; DB 162 DMLPEDTTHVRWFLNNKSVYSKDIKDIQIQGGQQLDLSTLNLNVTGTHSNYYSQS 221
; QY 167 GITVDLRLSKQAKFTVNFKNKTKKVIDL 195
; DB 222 AITDFE-KAPFGSKITV---DNTKNTIDV 246
;
; RESULT 9
; US-08-856-253-4
; Sequence 4, Application US/08856253
; Patent No. 6288214
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Patti, Joseph M.
; APPLICANT: House-Pompeo, Karen
; APPLICANT: Sthanam, Narayana
; APPLICANT: Symersky, Jindrich
; TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
;

```


QY 175 LSKQAKFTVFNKNGTKKVI 193
Db 605 ISGFDKAEITAKEGADLII 623

RESULT 12
US-09-206-942-43
; Sequence 43, Application US/09206942
; Patent No. 6432669
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High
; TITLE OF INVENTION: Molecular Weight Proteins
; FILE REFERENCE: 1038-861 MIS:jb
; CURRENT APPLICATION NUMBER: US/09/206,942
; CURRENT FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: 09/167,568
; EARLIER FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 1101
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-206-942-43

Query Match 10.8%; Score 110; DB 4; Length 1101;
Best Local Similarity 25.1%; Pred. No. 0.11;
Matches 50; Conservative 29; Mismatches 76; Indels 44; Gaps 8;

QY 18 ITVNGTSQNLSSLT-----FNKQOISYKDIENKVKSVLYFNKRGISDIDLRLSKQAKY 71
Db 452 LTVTGSAINIERKLTVEGSAKFLANPNYSFN-----VSLFDNQGKNSISI-----AKG 500
QY 72 TVHFK--NGTKRVVDLKAGIHTADLINTSD-----IKAISVNVDTKKQVKDKKAKANVQ 123
Db 501 GAHFKDINNTK-----SLNITNDSAVRTIIEGNIITNSGDLNITDNKNNAEIQ 550
QY 124 VPTYI-----TVNGTSQNLSSLTFFK--NQOISYKDLNENKVKSVLKNRGITDVLRL 174
Db 551 IGGNISQKEGNTLTSSDKINIITQITIKKGVNKGDSSTANNANLTIKTELQLTGDLN 610
QY 175 LSKQAKFTVFNKNGTKKVI 193
Db 611 ISGFDKAEITAKEGADLII 623

RESULT 13
US-09-206-942-28
; Sequence 28, Application US/09206942
; Patent No. 6432669
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High
; TITLE OF INVENTION: Molecular Weight Proteins
; FILE REFERENCE: 1038-861 MIS:jb
; CURRENT APPLICATION NUMBER: US/09/206,942
; CURRENT FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: 09/167,568
; EARLIER FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 1220
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-206-942-28

Query Match 10.8%; Score 110; DB 4; Length 1220;
Best Local Similarity 23.7%; Pred. No. 0.12;
Matches 54; Conservative 28; Mismatches 82; Indels 64; Gaps 9;

QY 3 GSHHHHHGSGIPTYI-----TVNGTSQNLSSLTFFNKNQOISYKDIENKVKSV 51
Db 422 GSDFDNH---QKPLTIKKDVIINSNLTAGGNVINGNLTAVNNGANL--KAITNFTENV 476
QY 52 --LYFNRGISDIDLRLSKQAKYTVHFKNKTRVVDLKAGIHTADLINTSDIKAISVNVDT 109
Db 477 GGLFDNKGNSISI-----ARGGAKFKDINNNTSSLN-ITNSTD 514
QY 110 KKQ-----VKDKKAKANVQVPTYI-----TVNGTSQNLSSLTFFK-- 145
Db 515 TYRTIIEGNIITKAGDLNIIDNKGNAEIQIGNISQKEGNTLTSSDKINIITQITIKKV 574
QY 146 NQOISYKDLNENKVKSVLKNRGITDVLRLSKQAKFTVFNKNGTKKVI 193
Db 575 NKEDSDSSTANNANLTIKTELQLTGDLNISGFDKAEITAKEGADLII 622

RESULT 14
US-09-206-942-26
; Sequence 26, Application US/09206942
; Patent No. 6432669
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High
; TITLE OF INVENTION: Molecular Weight Proteins
; FILE REFERENCE: 1038-861 MIS:jb
; CURRENT APPLICATION NUMBER: US/09/206,942
; CURRENT FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: 09/167,568
; EARLIER FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 1226
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-206-942-26

Query Match 10.8%; Score 110; DB 4; Length 1226;
Best Local Similarity 23.7%; Pred. No. 0.12;
Matches 54; Conservative 28; Mismatches 82; Indels 64; Gaps 9;

QY 3 GSHHHHHGSGIPTYI-----TVNGTSQNLSSLTFFNKNQOISYKDIENKVKSV 51
Db 428 GSDFDNH---QKPLTIKKDVIINSNLTAGGNVINGNLTAVNNGANL--KAITNFTENV 482
QY 52 --LYFNRGISDIDLRLSKQAKYTVHFKNKTRVVDLKAGIHTADLINTSDIKAISVNVDT 109
Db 483 GGLFDNKGNSISI-----ARGGAKFKDINNNTSSLN-ITNSTD 520
QY 110 KKQ-----VKDKKAKANVQVPTYI-----TVNGTSQNLSSLTFFK-- 145
Db 521 TYRTIIEGNIITKAGDLNIIDNKGNAEIQIGNISQKEGNTLTSSDKINIITQITIKKV 580
QY 146 NQOISYKDLNENKVKSVLKNRGITDVLRLSKQAKFTVFNKNGTKKVI 193
Db 581 NKEDSDSSTANNANLTIKTELQLTGDLNISGFDKAEITAKEGADLII 628

RESULT 15
US-09-134-001C-4005
; Sequence 4005, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

Db 457 -----SPKNERKKVLDLR 469

RESULT 21

US-09-023-082A-30

Sequence 30, Application US/09023082A

Patent No. 6077692

GENERAL INFORMATION:

APPLICANT: RUBEN, STEVEN M.

APPLICANT: JIMENEZ, PABLO

APPLICANT: DUAN, D. ROXANNE

APPLICANT: RAMPY, MARK A.

APPLICANT: MENDRICK, DONNA

APPLICANT: ZHANG, JUN

APPLICANT: NI, JIAN

APPLICANT: MOORE, PAUL A.

APPLICANT: COLEMAN, TIMOTHY A.

APPLICANT: GRUBER, JOACHIM R.

APPLICANT: DILLON, PATRICK J.

APPLICANT: GENTZ, REINER L.

TITLE OF INVENTION: KERATINOCYTE GROWTH FACTOR-2

NUMBER OF SEQUENCES: 148

CORRESPONDENCE ADDRESS:

ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.

STREET: 1100 NEW YORK AVE, NW, SUITE 600

CITY: WASHINGTON

STATE: DC

COUNTRY: USA

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/023,082A

FILING DATE: 13-FEB-1998

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/01790

FILING DATE: 14-FEB-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/461,195

FILING DATE: 05-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/023,852

FILING DATE: 13-AUG-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/039,045

FILING DATE: 28-FEB-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/862,432

FILING DATE: 23-MAY-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/910,875

FILING DATE: 13-AUG-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/055,561

FILING DATE: 13-AUG-1997

ATTORNEY/AGENT INFORMATION:

NAME: STEFFE, ERIC K.

REGISTRATION NUMBER: 36,688

REFERENCE/DOCKET NUMBER: 1488.0360008/EKS

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 184 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-023-082A-30

Query Match 10.1%; Score 102.5; DB 3; Length 184;

Best Local Similarity 24.3%; Pred. No. 0.043;

Matches 41; Conservative 26; Mismatches 61; Indels 41; Gaps 8;

Qy 1 MRGSHHHHGSQIPYITVNGTSQNILSSLTFFKNQOISYKDINKVKSVLYFN--RGI 58

Db 1 MRGSHHHHHS-----COALGDMVSPBATNSSSS-SFSSPSSAGRHRVRSYNHLOG- 51

Qy 59 SDIDLRLSKQAKYTVHF-----KN-----GTRK-----VVDLKAGIHTADLINTSDIK 101

Db 52 ---DVRWRKLPSPFKYFKIEKNGKVGSGTKKENCPCYSILEITSVEIGVAVKAINSNYYL 108

Qy 102 AIS-----VNVDTKKQVKDEKAKANQVQPYITVNGTSQNILSN 140

Db 109 AMNKKGLYGSKEFNNDCKLERIEENGYNITYASFNMQHNGRQMYVALN 157

RESULT 22

US-09-248-998-30

Sequence 30, Application US/09248998

Patent No. 6599879

GENERAL INFORMATION:

APPLICANT: Jimenez, Pablo

APPLICANT: Rampy, Mark A.

APPLICANT: Mendrick, Donna

APPLICANT: Russell, Deborah

APPLICANT: Louie, Arthur

TITLE OF INVENTION: Therapeutic Uses of Keratinocyte Growth Factor-2

FILE REFERENCE: 1488.1060002

CURRENT APPLICATION NUMBER: US/09/248,998

CURRENT FILING DATE: 1999-02-12

EARLIER APPLICATION NUMBER: US 60/114,387

EARLIER FILING DATE: 30-DEC-1998

EARLIER APPLICATION NUMBER: US 60/074,585

EARLIER FILING DATE: 13-FEB-1998

NUMBER OF SEQ ID NOS: 148

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 30

LENGTH: 184

TYPE: PRT

ORGANISM: Homo sapiens

US-09-248-998-30

Query Match 10.1%; Score 102.5; DB 4; Length 184;

Best Local Similarity 24.3%; Pred. No. 0.043;

Matches 41; Conservative 26; Mismatches 61; Indels 41; Gaps 8;

Qy 1 MRGSHHHHGSQIPYITVNGTSQNILSSLTFFKNQOISYKDINKVKSVLYFN--RGI 58

Db 1 MRGSHHHHHS-----COALGDMVSPBATNSSSS-SFSSPSSAGRHRVRSYNHLOG- 51

Qy 59 SDIDLRLSKQAKYTVHF-----KN-----GTRK-----VVDLKAGIHTADLINTSDIK 101

Db 52 ---DVRWRKLPSPFKYFKIEKNGKVGSGTKKENCPCYSILEITSVEIGVAVKAINSNYYL 108

Qy 102 AIS-----VNVDTKKQVKDEKAKANQVQPYITVNGTSQNILSN 140

Db 109 AMNKKGLYGSKEFNNDCKLERIEENGYNITYASFNMQHNGRQMYVALN 157

RESULT 23

US-09-610-651-30

Sequence 30, Application US/09610651

Patent No. 6693077

GENERAL INFORMATION:

APPLICANT: Ruben, Steven M.

APPLICANT: Jimenez, Pablo

APPLICANT: Duan, D. Roxanne

APPLICANT: Rampy, Mark A.

APPLICANT: Mendrick, Donna

APPLICANT: Zhang, Jun

APPLICANT: Ni, Jian

QY 3 GSHHHHGSQIPYTI-----TVNGTSQNLSSLTFNKNQOISYKDIENKVKSV 51
DB 422 GSDFDNH---QKPLTIKDVIIINSGLNTAGGIVNIAGNLTVESN--ANFKAITNFTFNV 476
QY 52 --LYFNRGISDI-----DLRLSKQAKYTVHFNGTKRVDLKGAGIHTADLIN-TS 98
DB 477 GGLFDNKGNSNLSIAKGGARFKDIDNSKLSITTNSSSTYRTII-----SGNITNKG 529
QY 99 DIKAISSVNDTKQ-----VKDKEAKANVQVPTITVNGTSQNLSSLTFNKNQOISYKDIENKVKSV 152
DB 530 DLNITNEGSDTEMQIGDVSKQEG-----NLTISSDKINITKQITIKAGVDGENSDS 581
QY 153 DLENNVKSVLKNSRGITVDLRLSKQAKFTVNFNGTKRVDLKGAGIHTADLIN-TS 198
DB 582 DATNANLTIKTKEKLTQDLNLSGFKAEITAKDGS-----DLTIG 623

RESULT 26
US-08-038-682-2
; Sequence 2, Application US/08038682
; Patent No. 5549897
; GENERAL INFORMATION:
; APPLICANT: BARENKAMP, STEPHEN J
; APPLICANT: ST. GEME III, JOSEPH W
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT SURFACE PROTEINS
; TITLE OF INVENTION: OF NON-TYPEABLE HAEMOPHILUS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/038,682
; FILING DATE: 16-MAR-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: BERKSTRESSER, JERRY W
; REGISTRATION NUMBER: 22,651
; REFERENCE/DOCKET NUMBER: 1038-293
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1536 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-038-682-2

Query Match 9.9%; Score 101; DB 1; Length 1536;
Best Local Similarity 23.9%; Pred. No. 1.2;
Matches 54; Conservative 32; Mismatches 86; Indels 54; Gaps 11;
QY 3 GSHHHHGSQIPYTI-----TVNGTSQNLSSLTFNKNQOISYKDIENKVKSV 51
DB 863 GSDFDNH---QKPLTIKDVIIINSGLNTAGGIVNIAGNLTVESN--ANFKAITNFTFNV 917
QY 52 --LYFNRGISDI-----DLRLSKQAKYTVHFNGTKRVDLKGAGIHTADLIN-TS 98
DB 918 GGLFDNKGNSNLSIAKGGARFKDIDNSKLSITTNSSSTYRTII-----SGNITNKG 970
QY 99 DIKAISSVNDTKQ-----VKDKEAKANVQVPTITVNGTSQNLSSLTFNKNQOISYKDIENKVKSV 152

DB 971 DLNITNEGSDTEMQIGDVSKQEG-----NLTISSDKINITKQITIKAGVDGENSDS 1022
QY 153 DLENNVKSVLKNSRGITVDLRLSKQAKFTVNFNGTKRVDLKGAGIHTADLIN-TS 198
DB 1023 DATNANLTIKTKEKLTQDLNLSGFKAEITAKDGS-----DLTIG 1064
RESULT 27
US-08-302-832-2
; Sequence 2, Application US/08302832
; Patent No. 5603938
; GENERAL INFORMATION:
; APPLICANT: BARENKAMP, STEPHEN J
; TITLE OF INVENTION: High Molecular Weight Surface Proteins
; TITLE OF INVENTION: of No. 5603938-Typeable Haemophilus
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd.
; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/302,832
; FILING DATE: 16-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9205704.1
; FILING DATE: 16-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US pct/us93/02166
; FILING DATE: 16-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Berkstresser, Jerry W
; REGISTRATION NUMBER: 22,651
; REFERENCE/DOCKET NUMBER: 1038-404
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1536 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-302-832-2
Query Match 9.9%; Score 101; DB 1; Length 1536;
Best Local Similarity 23.9%; Pred. No. 1.2;
Matches 54; Conservative 32; Mismatches 86; Indels 54; Gaps 11;
QY 3 GSHHHHGSQIPYTI-----TVNGTSQNLSSLTFNKNQOISYKDIENKVKSV 51
DB 863 GSDFDNH---QKPLTIKDVIIINSGLNTAGGIVNIAGNLTVESN--ANFKAITNFTFNV 917
QY 52 --LYFNRGISDI-----DLRLSKQAKYTVHFNGTKRVDLKGAGIHTADLIN-TS 98
DB 918 GGLFDNKGNSNLSIAKGGARFKDIDNSKLSITTNSSSTYRTII-----SGNITNKG 970
QY 99 DIKAISSVNDTKQ-----VKDKEAKANVQVPTITVNGTSQNLSSLTFNKNQOISYKDIENKVKSV 152
DB 971 DLNITNEGSDTEMQIGDVSKQEG-----NLTISSDKINITKQITIKAGVDGENSDS 1022
QY 153 DLENNVKSVLKNSRGITVDLRLSKQAKFTVNFNGTKRVDLKGAGIHTADLIN-TS 198

Db 1023 DATNNANLTIKTTELKLTQDLNLSGFNKAETAKDGS-----DLTIG 1064

RESULT 28

US-08-530-198-2
; Sequence 2, Application US/08530198

; Patent No. 5869065
; GENERAL INFORMATION:

; APPLICANT: BARENKAMP, STEPHEN J

; APPLICANT: ST. GENE III, JOSEPH W

; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT SURFACE PROTEINS

; TITLE OF INVENTION: OF NON-TYPEABLE HAEMOPHILUS

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Shoemaker and Mattare, Ltd

; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza

; STREET: Bldg. 1

; CITY: Arlington

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 22202-0286

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/530,198

; FILING DATE: 13-DEC-1995

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: BERKSTRESSER, JERRY W

; REGISTRATION NUMBER: 22,651

; REFERENCE/DOCKET NUMBER: JWB-1186

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 415-0810

; TELEFAX: (703) 415-0813

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1536 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-530-198-2

Query Match 9.9%; Score 101; DB 2; Length 1536;
Best Local Similarity 23.9%; Pred. No. 1.2;
Matches 54; Conservative 32; Mismatches 86; Indels 54; Gaps 11;

QY 3 GSHHHHHGSGIPIYTI-----TVNGTSQNLSSLTFTNKNQOISYKDIENKVKSV 51
Db 863 GSDFDNH---QKPLTIKDVINSGLTAGNIVNIAGNLTVESN--ANFKAITNFTNV 917

QY 52 --LYFNRGISDI-----DLRLSKQAKYTVHFKNGTKRVDLKGAIHTADLIN-TS 98
Db 918 GGLFDNKGNSISIAKGGARFKDIDNSKNLSITNSSSTYRII-----SGNITNKG 970

QY 99 DIKAISVNDTKQ-----VKDKEAKANVQVPYITVNGTSQNLNLTFFK--NQOISYK 152
Db 971 DLNITNEGSDTEMQIGDVSQKEG-----NLTSSDKINITKQITIKAGVDGENSDS 1022

QY 153 DLENNVKSVLKNSRGITDVLRLSKQAKFTVNFKNGTKKVIDLKAG 198
Db 1023 DATNNANLTIKTTELKLTQDLNLSGFNKAETAKDGS-----DLTIG 1064

RESULT 29

US-08-469-880-2
; Sequence 2, Application US/08469880

; Patent No. 5876733

; GENERAL INFORMATION:

; APPLICANT: BARENKAMP, STEPHEN J

; APPLICANT: Barenkamp, Stephen J.
; TITLE OF INVENTION: High Molecular Weight Surface Proteins
; TITLE OF INVENTION: OF No. 5876733-Typeable Haemophilus
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Shoemaker and Mattare, Ltd.

; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza

; STREET: Bldg. 1

; CITY: Arlington

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 22202-0286

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/469,880

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 9205704.1

; FILING DATE: 16-MAR-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US PCT/US93/02166

; FILING DATE: 16-MAR-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/302,832

; FILING DATE: 16-SEP-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Berkstresser, Jerry W

; REGISTRATION NUMBER: 22,651

; REFERENCE/DOCKET NUMBER: 1038-516 MIS-VG

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 415-0810

; TELEFAX: (703) 415-0813

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1536 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-08-469-880-2

Query Match 9.9%; Score 101; DB 2; Length 1536;
Best Local Similarity 23.9%; Pred. No. 1.2;
Matches 54; Conservative 32; Mismatches 86; Indels 54; Gaps 11;

QY 3 GSHHHHHGSGIPIYTI-----TVNGTSQNLSSLTFTNKNQOISYKDIENKVKSV 51
Db 863 GSDFDNH---QKPLTIKDVINSGLTAGNIVNIAGNLTVESN--ANFKAITNFTNV 917

QY 52 --LYFNRGISDI-----DLRLSKQAKYTVHFKNGTKRVDLKGAIHTADLIN-TS 98
Db 918 GGLFDNKGNSISIAKGGARFKDIDNSKNLSITNSSSTYRII-----SGNITNKG 970

QY 99 DIKAISVNDTKQ-----VKDKEAKANVQVPYITVNGTSQNLNLTFFK--NQOISYK 152
Db 971 DLNITNEGSDTEMQIGDVSQKEG-----NLTSSDKINITKQITIKAGVDGENSDS 1022

QY 153 DLENNVKSVLKNSRGITDVLRLSKQAKFTVNFKNGTKKVIDLKAG 198
Db 1023 DATNNANLTIKTTELKLTQDLNLSGFNKAETAKDGS-----DLTIG 1064

RESULT 30

US-08-470-2
; Sequence 2, Application US/08728470

; Patent No. 5928651

; GENERAL INFORMATION:

; APPLICANT: Barenkamp, Stephen J

;; TITLE OF INVENTION: High Molecular Weight Surface Proteins
;; NUMBER OF SEQUENCES: 10
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Shoemaker and Mattare, Ltd.
;; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
;; CITY: Arlington
;; STATE: Virginia
;; COUNTRY: U.S.A.
;; ZIP: 22202-0286
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/728,470
;; FILING DATE:
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/302,832
;; FILING DATE: 16-MAR-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US PCT/US93/02166
;; FILING DATE: 16-MAR-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: GB 9205704.1
;; FILING DATE: 16-MAR-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Berkstresser, Jerry W
;; REGISTRATION NUMBER: 22,651
;; REFERENCE/DOCKET NUMBER: 1038-633
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 415-0810
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1536 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE:
US-08-728-470-2

Query Match 9.9%; Score 101; DB 2; Length 1536;
Best Local Similarity 23.9%; Pred. No. 1.2;
Matches 54; Conservative 32; Mismatches 86; Indels 54; Gaps 11;
QY 3 GSHHHHGGSQIPYTI-----TVNGTSQNILSSTFNKNQOISYKDIENKVKSV 51
DB 863 GSDFDNH---QKPLTIKKDVIINSGLTAGNIVNIAGNLTVESN--ANFKAITNFTFNV 917
QY 52 --LYFNRGISDI-----DLRLSKQAKYTVHFKNKGTKRVDLKGAIHTADLIN-TS 98
DB 918 GGLFDNKGNSISIAKGGARPKDIDNSKNLSITTSNSTYRTII-----SGNITNKG 970
QY 99 DIKALSVNVDTKQ-----VKDKEAKANVQVPTITVNGTSQNILSNTLTFKK--NQOISYK 152
DB 971 DLNITNEGSDTEWQIGDVSQKEG-----NLTISSDKINITKQITIKAGVDGSDS 1022
QY 153 DLNNVKSVLKSNRGITDVLRLSKQAKFTVNFKNKGTKKVIDLKAG 198
DB 1023 DATNNANLTIKTKEKLQDLNLSGPNKAEITAKDGS-----DLTIG 1064

RESULT 31
US-08-617-697-2
; Sequence 2, Application US/08/617697
; Patent No. 5977336
; GENERAL INFORMATION:
; APPLICANT: Barenkamp, Stephen J
; TITLE OF INVENTION: High Molecular Weight Surface Proteins

;; TITLE OF INVENTION: of No. 5977336-Typeable Haemophilus
;; NUMBER OF SEQUENCES: 11
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Shoemaker and Mattare, Ltd.
;; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
;; CITY: Arlington
;; STATE: Virginia
;; COUNTRY: U.S.A.
;; ZIP: 22202-0286
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/617,697
;; FILING DATE: 01-APR-1996
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/302,832
;; FILING DATE: 05-OCT-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US PCT/US93/02166
;; FILING DATE: 16-MAR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Berkstresser, Jerry W
;; REGISTRATION NUMBER: 22,651
;; REFERENCE/DOCKET NUMBER: 1038-557
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 415-0810
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1536 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
US-08-617-697-2

Query Match 9.9%; Score 101; DB 2; Length 1536;
Best Local Similarity 23.9%; Pred. No. 1.2;
Matches 54; Conservative 32; Mismatches 86; Indels 54; Gaps 11;
QY 3 GSHHHHGGSQIPYTI-----TVNGTSQNILSSTFNKNQOISYKDIENKVKSV 51
DB 863 GSDFDNH---QKPLTIKKDVIINSGLTAGNIVNIAGNLTVESN--ANFKAITNFTFNV 917
QY 52 --LYFNRGISDI-----DLRLSKQAKYTVHFKNKGTKRVDLKGAIHTADLIN-TS 98
DB 918 GGLFDNKGNSISIAKGGARPKDIDNSKNLSITTSNSTYRTII-----SGNITNKG 970
QY 99 DIKALSVNVDTKQ-----VKDKEAKANVQVPTITVNGTSQNILSNTLTFKK--NQOISYK 152
DB 971 DLNITNEGSDTEWQIGDVSQKEG-----NLTISSDKINITKQITIKAGVDGSDS 1022
QY 153 DLNNVKSVLKSNRGITDVLRLSKQAKFTVNFKNKGTKKVIDLKAG 198
DB 1023 DATNNANLTIKTKEKLQDLNLSGPNKAEITAKDGS-----DLTIG 1064

RESULT 32
US-08-719-641-2
; Sequence 2, Application US/08/719641
; Patent No. 6218141
; GENERAL INFORMATION:
; APPLICANT: Barenkamp, Stephen J
; TITLE OF INVENTION: High Molecular Weight Surface Proteins
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd.

Wed Mar 16 14:08:18 2005

STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
STREET: Bldg. 1
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202-0286
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/719,641
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/302,832
FILING DATE: 16-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US93/02166
FILING DATE: 16-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9205704.1
FILING DATE: 16-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Berkstreiser, Jerry W
REGISTRATION NUMBER: 22,651
REFERENCE/DOCKET NUMBER: 1038-625
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 415-0810
TELEFAX: (703) 415-0813
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1536 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
US-08-719-641-2

Query Match 9.9%; Score 101; DB 3; Length 1536;
Best Local Similarity 23.9%; Pred. No. 1.2;
Matches 54; Conservative 32; Mismatches 86; Indels 54; Gaps 11;
QY 3 GSHHHHGSQIPYTI-----TVNGTSQNLSSLTFFNKQOISYKDIEKVKSV 51
DB 863 GSDFDNH---QKPLTIKQVINSGLTAGNIVNAGNLTVESN--ANFKAITNFTFV 917
QY 52 --LYFNRGISDI-----DLRLSKQAKYTVHFKNKTKRVVDLKGAGHTADLIN-TS 98
DB 918 GGLFDNKGNSNISIAGGARFKDIDNSKLSITNSSSTYRTII-----SGNITKNG 970
QY 99 DIKAISVNVDTKKQ-----VKDKEAKANVQVPYITVNGTSQNLSSLTFFNK--NQOISYK 152
DB 971 DLNITNEGSDTEMQIGDVSQKEG-----NLTISSDKINIKTKQITIKAGVDGENSDS 1022
QY 153 DLENNVSKVLKSNRGITDVLRLSKQAKYTVNFKNGTKKVIDLKAG 198
DB 1023 DATNNANLTIKTELKLTQDLNLSGFKNAEITAKDGS----DLTIG 1064

RESULT 33
US-09-206-942-67
Sequence 67, Application US/09206942
Patent No. 6432669
GENERAL INFORMATION:
APPLICANT: Loosmore, Sheena M.
APPLICANT: Yang, Yan-ping
APPLICANT: Klein, Michel H.
TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High
FILE REFERENCE: 1038-861 MIS:jb
CURRENT APPLICATION NUMBER: US/09/206,942

CURRENT FILING DATE: 1998-12-08
EARLIER APPLICATION NUMBER: 09/167,568
EARLIER FILING DATE: 1998-10-07
NUMBER OF SEQ ID NOS: 95
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 67
LENGTH: 1536
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-09-206-942-67
Query Match 9.9%; Score 101; DB 4; Length 1536;
Best Local Similarity 23.9%; Pred. No. 1.2;
Matches 54; Conservative 32; Mismatches 86; Indels 54; Gaps 11;
QY 3 GSHHHHGSQIPYTI-----TVNGTSQNLSSLTFFNKQOISYKDIEKVKSV 51
DB 863 GSDFDNH---QKPLTIKQVINSGLTAGNIVNAGNLTVESN--ANFKAITNFTFV 917
QY 52 --LYFNRGISDI-----DLRLSKQAKYTVHFKNKTKRVVDLKGAGHTADLIN-TS 98
DB 918 GGLFDNKGNSNISIAGGARFKDIDNSKLSITNSSSTYRTII-----SGNITKNG 970
QY 99 DIKAISVNVDTKKQ-----VKDKEAKANVQVPYITVNGTSQNLSSLTFFNK--NQOISYK 152
DB 971 DLNITNEGSDTEMQIGDVSQKEG-----NLTISSDKINIKTKQITIKAGVDGENSDS 1022
QY 153 DLENNVSKVLKSNRGITDVLRLSKQAKYTVNFKNGTKKVIDLKAG 198
DB 1023 DATNNANLTIKTELKLTQDLNLSGFKNAEITAKDGS----DLTIG 1064

RESULT 34
US-09-976-594-4
Sequence 4, Application US/09976594
Patent No. 6673549
GENERAL INFORMATION:
APPLICANT: Furness, Michael
APPLICANT: Buchbinder, Jenny
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
FILE REFERENCE: PA-0041 US
CURRENT APPLICATION NUMBER: US/09/976,594
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: 60/240,409
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 1143
SOFTWARE: PERL Program
SEQ ID NO 4
LENGTH: 1207
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6673549 1863336CD1
US-09-976-594-4

Query Match 9.8%; Score 99.5; DB 4; Length 1207;
Best Local Similarity 24.0%; Pred. No. 1.2;
Matches 48; Conservative 36; Mismatches 79; Indels 37; Gaps 8;
QY 21 NGTSQNLSSLTFFNKQOISYKDIEKVKSVLYFNRG-----SDIDLRLSKQAKYTVHF 75
DB 331 NGNLDVLSFLKSN-TLPISLQSTRSSL--LLGFRSMNININLWSITELLYKGRYCOEP 387
QY 76 KNGTKRVVDLKGAGHTADLIN-TS DIKAISVNVDTKKQVKDKEAKANVQVPYITVNGTSQ 135
DB 388 RGPTAV-----GYFLYNLIDMSSEVQAKEHLRQYFHOLEKKNVPIENI-YRGI-R 440
QY 136 NTL-----SNLTFKNQOISYKDLNNSVSKNKGITDVLRLSK 177
DB 441 NLLESVHVPELTKDAHLAVESKNLDFQKTVQTSSELESTLETLEAKNQPIRDV-----LK 496
QY 178 QAKFTVNFKNKTKKVIDLKA 197

[illegible]

US-09-995-587A-1

Query Match 9.5%; Score 96.5; DB 4; Length 789;
Best Local Similarity 24.9%; Pred. No. 1.3;
Matches 48; Conservative 31; Mismatches 87; Indels 27; Gaps 8;

QY 19 TVNGTSQLSLTFNKNQOISYKDIENKVKSVLYFNRI---SDIDLRLS---KQKYYT 72
DB 24 TVNASADTNIEN-NDSSVTQVTTGNDIAVKSVTILGSGQVSAASDPTTIRTSANANSASSA 82

QY 73 VHFKNKTKRVVD---LKAGIHTADLINTSDIKALSIVNVD--KKQVKDEAKAN---VQV 124
DB 83 ANTONSNSQVASSAAITSTSSAASLNNTDSKAAQENTNTAKNDTQKAPANESSAKN 142

QY 125 PYTIVNGTS-----QNILSNITFFKKNQOISYKDLNNVSVLKNRGITVDVLRSLK 177
DB 143 EPAVNVDSSAAKNDQSSKNNTAKLN-----KDAENVVKKAGIDPNSLTDDQIKALN 197

QY 178 QAKFTVFNFKNGTK 190
DB 198 KMFPSKAAKSGTQ 210

RESULT 39
US-08-621-081A-19
; Sequence 19, Application US/08621081A
; Patent No. 5795974
; GENERAL INFORMATION:
; APPLICANT: Cole, Barry C.
; APPLICANT: Atkin, Curtis L.
; APPLICANT: Knudsen, Kevin L.
; APPLICANT: Sawitzke, Allen D.
; TITLE OF INVENTION: Mycoplasma Arthritis Superantigen
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thorpe, No. 5795974th & Western
; STREET: 9035 South 700 East, Suite 200
; CITY: Sandy
; STATE: Utah
; COUNTRY: USA
; ZIP: 84070
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM ThinkPad 340
; OPERATING SYSTEM: MS-DOS 6.22
; SOFTWARE: Word Perfect 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/621,081A
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/165,038
; FILING DATE: December 10, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Alan J. Howarth
; REGISTRATION NUMBER: 36,553
; REFERENCE/DOCKET NUMBER: T676.CIP/U-1195
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (801)566-6633
; TELEFAX: (801)566-0750
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 221 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-621-081A-19

Query Match 9.4%; Score 95.5; DB 1; Length 221;
Best Local Similarity 20.9%; Pred. No. 0.26;
Matches 45; Conservative 35; Mismatches 64; Indels 71; Gaps 9;

QY 3 GSHHHHGGSQIPYITVNGTSQNLSTFNKNQOISYKDIENKVKSVLYFNRGISDID 62

DB 2 GHHHHHMKLRV-----ENPKKQAKHFQNLNLLNVVFTNKELEDI- 40

QY 63 LRLSKQAKYTVHFKNGTKRVW-----DLKAGIHTADLINTSDIKAI 103

DB 41 -----YNLSNKEETKEVLKLFKLKVNQVYRHAFGIVNDYNGLLLEYKEIFNMFLK-L 91

QY 104 SVNVDTKKQVKDEAKANVQVPTITVNGTSQNLSTFNKNQOISYKDLNNVSVLKN 163

DB 92 SVVFDIQR-----KEANNVQIKENIA-----ILDEIMAKANDLSYFISQN-----K 134

QY 164 SNRGITVDLRLSKQAKFTVFNFKNGTKKVIDLKAG 198

DB 135 NFQELMDKAVKLTKEMKIKL---KGQK--LDLRDG 164

RESULT 40
US-09-107-433-3308
; Sequence 3308, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3308:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...331
; SEQUENCE DESCRIPTION: SEQ ID NO: 3308:
US-09-107-433-3308

Query Match 9.4%; Score 95; DB 4; Length 331;
Best Local Similarity 22.9%; Pred. No. 0.51;
Matches 54; Conservative 32; Mismatches 64; Indels 86; Gaps 13;

QY 3 GSHHHHGGSQIPYITVNGTSQNLSTFNKNQOISYKDIENKVKSVLYFNRGISDID 62

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 16, 2005, 10:36:37 ; Search time 168 Seconds
(without alignments)
460.429 Million cell updates/sec

Title: US-10-041-775-2

Perfect score: 1016

Sequence: 1 MRGSHHHHHSQIPYITV.....FTVNFKNGTKKVIDLKAGIY 200

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003s:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1016	100.0	200	6	ABP58754 Recombina
2	950	93.5	689	2	AAW26301 Staphyloc
3	886.5	87.3	657	6	AAO27181 Staphyloc
4	818.5	80.6	584	6	ABM71428 Staphyloc
5	626	61.6	131	5	AAE25524 Staphyloc
6	626	61.6	131	6	ABP58755 Recombina
7	440	43.3	150	2	AAW89790 Staphyloc
8	210	20.7	144	6	ABM73216 Staphyloc
9	210	20.7	151	2	AAW89785 Staphyloc
10	159	15.6	31	2	AAW26302 Staphyloc
11	155	15.3	31	2	AAW26303 Staphyloc
12	142	14.0	141	6	ABJ19003 Pathogen
13	142	14.0	141	6	ABM73460 Staphyloc
14	135	13.3	96	6	ABM71429 Staphyloc
15	133	13.1	343	7	ABM79016 Staphyloc
16	131	12.9	316	7	ABM79017 Staphyloc
17	117.5	11.6	391	5	ABM79627 Plasmodi
18	117.5	11.6	391	6	ABP71275 E. coli e
19	117.5	11.6	391	2	ADM86524 E. coli e
20	116.5	11.5	512	2	AAW31554 Collagen
21	115	11.3	374	6	ABU44119 Protein e
22	114.5	11.3	393	5	ABB79625 E. coli e
23	114.5	11.3	393	6	ABP71273 P. falcip
24	114.5	11.3	393	7	ADM86520 pET42A Pl
25	114	11.2	211	2	AAW31553 Collagen

26	113.5	11.2	246	6	ABU25692 Protein e
27	111	10.9	963	6	ABU33173 Protein e
28	110	10.8	31	2	AAW26306 Staphyloc
29	110	10.8	438	8	ADG32048 Mutant B
30	110	10.8	1095	3	AAAB01835 Haemophil
31	110	10.8	1101	3	AAAB01834 Haemophil
32	110	10.8	1221	3	AAAB01825 Haemophil
33	110	10.8	1227	3	AAAB01824 Haemophil
34	108.5	10.7	288	5	ABP39160 Staphyloc
35	108.5	10.7	288	8	ADP39160 Staphyloc
36	108.5	10.7	371	7	ADP39160 Staphyloc
37	107.5	10.6	493	2	AAE24164 Plasmodi
38	106	10.4	345	2	AAE24164 Plasmodi
39	105	10.3	560	7	AAW31555 Fibronect
40	105	10.3	1781	8	ABM79015 Staphyloc
41	104	10.2	74	4	ADP25447 Plasmodi
42	104	10.2	159	2	AAW31552 Recogin
43	104	10.2	1051	5	ABP40160 Staphyloc
44	104	10.2	1051	6	ABU42634 Protein e
45	104	10.2	1051	8	ADS07207 Staphyloc

ALIGNMENTS

RESULT 1

ABP58754
ID ABP58754 standard; protein; 200 AA.
AC ABP58754;
XX
DT 03-APR-2003 (first entry)
DE Recombinant Staphylococcus aureus Map19.
XX

KW Map19; Map; major histocompatibility complex class II analogue protein;
KW recombinant; T-cell mediated response; prevention; modulation;
KW T-cell overstimulation; toxic shock syndrome; poison ivy;
KW T-cell lymphoproliferative disease; leukaemia; autoimmune disease;
KW delayed-type hypersensitivity response; DTH; staphylococcal infection;
KW immunomodulator; immunosuppressive; cytostatic; vaccine.
XX

OS Staphylococcus aureus.

OS Synthetic.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC overstimulation of T-cells such as toxic shock syndrome or poison ivy,
CC and T-cell lymphoproliferative diseases such as leukaemia and autoimmune
CC disease. They may also be used to reduce delayed-type hypersensitivity
CC (DTH) responses and to prevent or modulate a T-cell mediated response to
CC a staphylococcal infection. The present sequence represents recombinant
CC Map19
XX
SQ Sequence 200 AA;
Query Match 100.0%; Score 1016; DB 6; Length 200;
Best Local Similarity 100.0%; Pred. No. 2.8e-80;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRGSHHHHSGSQIPYITVNGTSQNLSSLTFFKNQOISYKDIENKVSQVLYFNRGISD 60
DB 1 MRGSHHHHSGSQIPYITVNGTSQNLSSLTFFKNQOISYKDIENKVSQVLYFNRGISD 60
QY 61 IDRLRSQAKYTVHFKNQTKRVVDLKGAGHTADLINTSDIKAISVNVDTKKQVCKEAKA 120
DB 61 IDRLRSQAKYTVHFKNQTKRVVDLKGAGHTADLINTSDIKAISVNVDTKKQVCKEAKA 120
QY 121 NVQVPYITVNGTSQNLSSLTFFKNQOISYKDIENKVSQVLYFNRGISD 180
DB 121 NVQVPYITVNGTSQNLSSLTFFKNQOISYKDIENKVSQVLYFNRGISD 180
QY 181 FTVNFKNGTKKVIDLKAGIY 200
DB 181 FTVNFKNGTKKVIDLKAGIY 200

RESULT 2
AAW26301
ID AAW26301 standard; protein; 689 AA.
XX
AC AAW26301;
XX
DT 17-OCT-2003 (revised)
DT 16-NOV-1997 (first entry)
XX
DE Staphylococcal MHC class II antigen analogue protein.
XX
KW MHC II analogue protein; major histocompatibility complex;
KW staphylococcal; adhesin; virulence factor; vaccine.
XX
OS Staphylococcus aureus; strain FDA 574.
XX
FH Key Location/Qualifiers
FT Peptide 1..30
FT Protein /label= Sig_peptide
FT /label= Mat_protein
XX
XX US5648240-A.
XX
XX
XX
PD 15-JUL-1997.
XX
XX 24-MAY-1994; 94US-00248021.
XX
XX 24-MAY-1994; 94US-00248021.
XX
XX (TEXA) UNIV TEXAS A & M.
XX
XX Jonsson K, Hook M, Patti JM, Gurusiddappa S;
XX WPI; 1997-372059/34.
XX N-PSDB; AAT84441.
XX
FT DNA encoding Staphylococcus aureus broad spectrum adhesin - for
FT production of recombinant adhesin for use in vaccines.
XX
PS Claim 17; Col 11-14; 30pp; English.
XX
XX This protein sequence comprises a broad spectrum adhesin of

CC Staphylococcus aureus FDA 574 that is capable of binding fibronectin or
CC vitronectin and which includes six repeated segments containing MHC II
CC mincing units (see AAW26302-07) of about 30 amino acid residues.
CC Studies of the binding specificities of the adhesin show that it
CC functionally resembles an MHC II antigen in that it binds synthetic
CC peptides. Vectors incorporating DNA (see AAT84441) encoding the adhesin
CC can be used to express recombinant adhesin in transformed host (esp.
CC bacterial, partic. E. coli) cells. Recombinant adhesin can be used in
CC vaccines against staphylococcal infections. (Updated on 17-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 689 AA;
Query Match 93.5%; Score 950; DB 2; Length 689;
Best Local Similarity 97.0%; Pred. No. 8.3e-74;
Matches 191; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 4 SHHHHHHSGSQIPYITVNGTSQNLSSLTFFKNQOISYKDIENKVSQVLYFNRGISD 63
DB 41 SLHHGYSKIQIPYITVNGTSQNLSSLTFFKNQOISYKDIENKVSQVLYFNRGISD 100
QY 64 RLSQAKYTVHFKNQTKRVVDLKGAGHTADLINTSDIKAISVNVDTKKQVCKEAKANVQ 123
DB 101 RLSQAKYTVHFKNQTKRVVDLKGAGHTADLINTSDIKAISVNVDTKKQVCKEAKANVQ 160
QY 124 VPYITVNGTSQNLSSLTFFKNQOISYKDIENKVSQVLYFNRGISD 183
DB 161 VPYITVNGTSQNLSSLTFFKNQOISYKDIENKVSQVLYFNRGISD 220
QY 184 NFKNGTKKVIDLKAGIY 200
DB 221 NFKNGTKKVIDLKAGIY 237
RESULT 3
AAO27181
ID AAO27181 standard; protein; 657 AA.
XX
AC AAO27181;
XX
DT 17-SEP-2003 (first entry)
XX
DE Staphylococcus aureus p70 protein, member of the EAP family.
XX
KW EAP; extracellular adherence protein; antiinflammatory; fibrinogen;
KW ICAM-1; radiation; allergin; sunburn; asthma; psoriasis; encephalitis;
KW meningitis; osteomyelitis; gingivitis; pancreatitis; antiadhesive;
KW cancer therapy; atherosclerosis; rheumatoid arthritis; strain wood 46.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT Region 22..50
FT /note= "Repeat region 1"
FT Region 131..159
FT /note= "Repeat region 1"
FT Region 332..362
FT /note= "Repeat region 2"
FT Region 437..477
FT /note= "Repeat region 2 with a partial 10 residue overlap
FT of repeat region 1"
FT Region 542..572
FT /note= "Repeat region 2"
XX
XX WO2003041726-A1.
XX
XX 22-MAY-2003.
XX
XX 14-NOV-2002; 2002WO-SR002075.
XX
XX 16-NOV-2001; 2001SE-00003831.
XX 16-NOV-2001; 2001US-0331456P.
XX 21-NOV-2001; 2001US-0331782P.

XX (BIOS-) BIOTAPRO AB.
XX Flock J, Herrmann M, Preissner KT, Chavakis T;
XX WPI; 2003-468585/44.
XX
XX Use of extracellular adherence protein for the manufacture of anti-
XX inflammatory or anticancer drug in the treatment of e.g. atherosclerosis.
XX
XX Disclosure; Page 6; 22pp; English.
XX
XX This invention relates to the use of extracellular adherence proteins
XX (EAPs) are used for antiinflammatory purposes to treat acute or chronic
XX inflammation. Proteins of the EAP family are produced by the bacterium
XX Staphylococcus aureus (S. aureus) and are able to bind to S. aureus cells
XX and agglutinate them. This invention, however, relates to the ability of
XX EAP to bind adhesive proteins in the connective tissue and on cell
XX surfaces, such that it blocks the inflammatory response of the infected
XX host organism. Specifically EAP can bind fibrinogen in the extracellular
XX matrix or ICAM-1 on cells to inhibit the mobility, infiltration and
XX activities of granulocytes, macrophages and lymphocytes. Note that the
XX inflammatory disorders can be from either a bacterial or non-bacterial
XX source and could include responses to radiation, infection, chemicals,
XX allergens and injury. As such EAP can be used as an antiinflammatory to
XX treat various conditions including sunburn, asthma, allergy, psoriasis,
XX encephalitis, meningitis, osteomyelitis, gingivitis and pancreatitis.
XX Furthermore, the antiadhesive potential of EAP provides a method for
XX cancer therapy and treatment of atherosclerosis and rheumatoid arthritis
XX caused by uncontrolled cellular extravasation of leukocytes. This
XX polypeptide sequence is the Staphylococcus aureus strain Wood 46 p70
XX protein, which is a typical member of the EAP family of the invention
XX
SQ Sequence 657 AA;

Query Match 87.3%; Score 886.5; DB 6; Length 657;
Best Local Similarity 90.8%; Pred. No. 2.6e-68;
Matches 177; Conservative 8; Mismatches 9; Indels 1; Gaps 1;

Qy 6 HHHHGSQIPYITVNGTSQNLSTFNKNQISYKDIENKVSVLVFNRGISDIDL 65
Db 13 HGHNSQTPIYITVNGTSQNLSTFNKNQISYKDIENKVSVLVFNRGISDIDL 72

Qy 66 SKQAKYTVHFKNQTKRVVDLKGAIHTADLINTSDIKAISVNVDTKKQVKDKKAKANQVP 125
Db 73 SKQAEYTVHFKNQTKRVVDLKGAIHTADLINTSDIKAISVNVDTKKQVKDKKAKANQVP 131

Qy 126 YTITVNGTSQNLSTFNKNQISYKDIENKVSVLVFNRGISDIDLRLSKQAKFTVNF 185
Db 132 YTITVNGTSQNLSTFNKNQISYKDIENKVSVLVFNRGISDIDLRLSKQAKFTVNF 191

Qy 186 KNGTKKKVIDLKAGIY 200
Db 192 KNGTKKKVIDLKAGIY 206

RESULT 4
ABM71428
ID ABM71428 standard; protein; 584 AA.
XX
XX AC ABM71428;
XX
XX DT 20-NOV-2003 (first entry)
XX
XX DE Staphylococcus aureus protein #669.
XX
XX KW Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
XX enzymatic assay; antibiotic target.
XX
XX OS Staphylococcus aureus.
XX
XX PN WO200294868-A2.
XX

PD 28-NOV-2002.
XX
XX PF 27-MAR-2002; 2002WO-IB002637.
XX
XX PR 27-MAR-2001; 2001GB-00007661.
XX
XX PA (CHIR-) CHIRON SPA.
XX
XX PI Masignani V, Mora M, Scarselli M;
XX
XX WPI; 2003-120786/11.
XX
XX DR N-PSDB; ACF72988.
XX
XX PT New Staphylococcus aureus protein, useful as a vaccine for treating or
XX preventing Staphylococcal infection, specifically an infection caused by
XX S. aureus, e.g. sepsis.
XX
XX PS Claim 1; SEQ ID NO 1336; 49pp; English.
XX
XX CC The invention relates to novel genes and encoded proteins from
XX Staphylococcus aureus. A composition comprising the S. aureus protein, a
XX nucleic acid encoding the protein, or an antibody to the protein, is
XX useful as a pharmaceutical, particularly as a vaccine for treating or
XX preventing infection due to Staphylococcus bacteria, specifically an
XX infection caused by S. aureus. The composition is particularly useful for
XX treating or preventing sepsis in a patient. The composition can also be
XX used for diagnostics. The protein is also used in an assay for enzymatic
XX studies and as a target for antibiotics. This sequence represents one of
XX the novel S. aureus proteins of the invention
XX
SQ Sequence 584 AA;

Query Match 80.6%; Score 818.5; DB 6; Length 584;
Best Local Similarity 82.3%; Pred. No. 1.8e-62;
Matches 163; Conservative 19; Mismatches 15; Indels 1; Gaps 1;

Qy 4 SHHHHGSQIPYITVNGTSQNLSTFNKNQISYKDIENKVSVLVFNRGISDIDL 63
Db 41 SLHHGYSKVHVPYAITVNGTSQNLSTFNKNQISYKDIENKVSVLVFNRGISDIDL 100

Qy 64 RLSQAKYTVHFKNQTKRVVDLKGAIHTADLINTSDIKAISVNVDTKKQVKDKKAKANV 122
Db 101 RLSQAKYTVHFKNQTKRVVDLKGAIHTADLINTSEIKAININVDTKKQVEDKKDKKANY 160

Qy 123 QVPYITVNGTSQNLSTFNKNQISYKDIENKVSVLVFNRGISDIDLRLSKQAKFT 182
Db 161 QVPYITVNGTSQNLSTFNKNQISYKDIENKVSVLVFNRGISDIDLRLSKQAKFT 220

Qy 183 VNFKNQTKKKVIDLKAGIY 200
Db 221 VNFKNQTKKKVIDLKAGIY 238

RESULT 5
AAE25524
ID AAE25524 standard; protein; 131 AA.
XX
XX AC AAE25524;
XX
XX DT 04-NOV-2002 (first entry)
XX
XX DE Staphylococcus aureus Map10 protein.
XX
XX KW Map10 protein; infection; immunological response; passive immunisation;
XX vaccine; antibacterial.
XX
XX OS Staphylococcus aureus.
XX
XX PN WO200234788-A1.
XX
XX PD 02-MAY-2002.
XX
XX PF 22-OCT-2001; 2001WO-US032550.
XX

XX 20-OCT-2000; 2000US-0241832P.
 PR 21-MAR-2001; 2001US-0277287P.
 XX (INHI-) INHIBITEX INC.
 PA Patti JM, Domanski P, Patel P;
 XX WPI; 2002-547435/58.
 XX N-PSDB; AAD41844.
 XX Antibody capable of binding subdomains of Map10 protein, a surface
 PT localized protein from Staphylococcus aureus, useful for preventing and
 PT treating staphylococcal infection in humans or animals.
 XX Claim 28; Page 46-47; 52pp; English.
 XX The present sequence relates to novel antibodies which are capable of
 CC binding to the subdomains Map10 protein, a surface localised protein from
 CC Staphylococcus aureus. Sequences of the invention prevent S.aureus
 CC infection in a human or animal and inhibit binding of staphylococcal
 CC bacteria to eukaryotic cells. They are also useful for diagnosing an
 CC infection of S.aureus in a sample and for treating or preventing S.aureus
 CC infection in humans or animals. Antibodies of the invention are useful
 CC for inducing an immunological response in a human or animal and for
 CC identifying antibodies to the Map10 protein in a sample. They are also
 CC useful in preventing adherence of Staphylococcal bacteria, in production
 CC facilities or laboratories to isolate additional quantities of the
 CC proteins, such as by affinity chromatography and in the development of
 CC vaccines for passive immunisation against staphylococcal infections. The
 CC antibodies prevent or reduce bacterial infection on in-dwelling medical
 CC devices to make them safer to use, including sutures, replacement heart
 CC valves, cardiac assist devices, hard and soft contact lenses, intraocular
 CC lens implants, other implants such as corneal inlays, kerato-prostheses,
 CC vascular stents, dental prostheses, pacemakers and heart valves. The
 CC present sequence is Staphylococcus aureus Map10 protein
 XX
 XX Sequence 131 AA;

Query Match 61.6%; Score 626; DB 5; Length 131;
 Best Local Similarity 100.0%; Pred. No. 1.3e-46;
 Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRGSHHHHGGSQIPYITVNGTSQNLSSLTFFNKNQIISYKDIEKVKSVLYFNRGISD 60
 DB 1 MRGSHHHHGGSQIPYITVNGTSQNLSSLTFFNKNQIISYKDIEKVKSVLYFNRGISD 60
 QY 61 IDRLSKQAKYTVHFKNKGTFRVVDLKAGIHTADLINTSDIKAISVNVDTKKQVKDKEAKA 120
 DB 61 IDRLSKQAKYTVHFKNKGTFRVVDLKAGIHTADLINTSDIKAISVNVDTKKQVKDKEAKA 120
 QY 121 NV 122
 DB 121 NV 122

RESULT 6
 ABP58755
 ID ABP58755 standard; protein; 131 AA.
 XX AC ABP58755;
 XX 03-APR-2003 (first entry)
 DT Recombinant Staphylococcus aureus Map10.
 XX Map10; Map; major histocompatibility complex class II analogue protein;
 KW recombinant; T-cell mediated response; prevention; modulation;
 KW T-cell overstimulation; toxic shock syndrome; poison ivy;
 KW T-cell lymphoproliferative disease; leukaemia; autoimmune disease;
 KW delayed-type hypersensitivity response; DTH; staphylococcal infection;
 KW immunomodulator; immunosuppressive; cytostatic; vaccine.
 XX

OS Staphylococcus aureus.
 OS Synthetic.
 XX WO200277010-A2.
 XX PD 03-OCT-2002.
 XX 10-JAN-2002; 2002WO-US000401.
 PF 10-JAN-2001; 2001US-0260523P.
 PR (TEXA) UNIV TEXAS A & M SYSTEM.
 XX Brown EN, Lee LY, Hook M;
 XX WPI; 2003-046760/04.
 DR N-PSDB; ABZ26000.
 XX New method of preventing or modulating T-cell-mediated response in a
 PT host, useful for e.g. the treatment of toxic shock syndrome, comprises
 PT administering a Staphylococcus aureus major histocompatibility complex
 PT class II analog protein.
 XX Claim 13; Page 55; 55pp; English.
 XX The invention relates to a novel method of preventing or modulating a T-
 CC cell mediated response in an individual involving the administration of
 CC the Staphylococcus aureus major histocompatibility complex class II
 CC analogue protein (Map), or its active fragments such as Map19 (ABP58754)
 CC and Map10 (ABP58755). Map, and its active fragments Map19 and Map10, may
 CC be used to treat or prevent pathogenic conditions associated with
 CC overstimulation of T-cells such as toxic shock syndrome or poison ivy,
 CC and T-cell lymphoproliferative diseases such as leukaemia and autoimmune
 CC disease. They may also be used to reduce delayed-type hypersensitivity
 CC (DTH) responses and to prevent or modulate a T-cell mediated response to
 CC a staphylococcal infection. The present sequence represents recombinant
 CC Map10
 XX
 XX Sequence 131 AA;

Query Match 61.6%; Score 626; DB 6; Length 131;
 Best Local Similarity 100.0%; Pred. No. 1.3e-46;
 Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRGSHHHHGGSQIPYITVNGTSQNLSSLTFFNKNQIISYKDIEKVKSVLYFNRGISD 60
 DB 1 MRGSHHHHGGSQIPYITVNGTSQNLSSLTFFNKNQIISYKDIEKVKSVLYFNRGISD 60
 QY 61 IDRLSKQAKYTVHFKNKGTFRVVDLKAGIHTADLINTSDIKAISVNVDTKKQVKDKEAKA 120
 DB 61 IDRLSKQAKYTVHFKNKGTFRVVDLKAGIHTADLINTSDIKAISVNVDTKKQVKDKEAKA 120
 QY 121 NV 122
 DB 121 NV 122

RESULT 7
 AAW89790
 ID AAW89790 standard; protein; 150 AA.
 XX AC AAW89790;
 XX 16-MAR-1999 (first entry)
 DT Staphylococcus aureus protein SEQ ID #5238.
 DE Computer readable medium; vaccine; S.aureus infection; immunodetection;
 XX cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KW skin infection; surgical wound infection; scalded skin syndrome;
 KW toxic shock syndrome.
 XX Staphylococcus aureus.
 OS

XX Key Location/Qualifiers
FH Peptide 1. .12
FT /note= "vector pQE30-derived peptide"
FT 13. .512
FT /note= "epitope M17"
XX
XX WO9743314-A2.
PN
XX 20-NOV-1997.
PD
XX 14-MAY-1997; 97WO-US008210.
PF
XX 16-MAY-1996; 96US-0017678P.
PR
XX (TEXA) UNIV TEXAS A & M SYSTEM.
PA (UABR-) UAB RES FOUND.
XX
XX Hoecek M, Patti JM, House-Pompeo K, Schanam N, Symersky J;
PI
XX WPI; 1998-008801/01.
DR N-PSDB; AAT93438.
XX
XX Antibody that interacts with collagen binding domain of Staphylococcal
PT cna gene product - useful to prevent bacterial sepsis in animal infected
PT with Staphylococcus aureus.
XX
XX Claim 31; Page 117-119; 143pp; English.
XX
XX This protein comprises Staphylococcus aureus collagen binding protein
CC (CBP) epitope M55, i.e. amino acids 30-531 of full-length CBP, plus a
CC vector-derived N-terminal peptide. Claimed 441, 849 and 1500 bp nucleic
CC acid sequences (see AAT93436-38) respectively encode CBP epitopes M17,
CC M31 and M55 (see AAW31552-54) that confer protection against S. aureus
CC infection. These nucleic acid sequences can be used in the recombinant
CC production of the CBP epitopes. The CBP protein and antigenic epitopes
CC are contemplated for use in the treatment of pathological infections,
CC especially to prevent bacterial adhesion to collagen. The claimed nucleic
CC acids as well as claimed anti-CBP antibodies will also be of use in
CC screening, diagnostic and therapeutic applications including active and
CC passive immunisation and methods for the prevention of bacterial
CC colonisation in an animal such as a human. The CBP epitopes are also
CC contemplated for use in the preparation of vaccines and as carrier
CC proteins in vaccine formulations, as well as in the formulation of
CC compositions for the prevention of S. aureus infection. (Updated on 25-
CC MAR-2003 to correct PI field.) (Updated on 27-AUG-2003 to correct OS
CC field.)
XX Sequence 512 AA;
SQ
Query Match 11.5%; Score 116.5; DB 2; Length 512;
Beat Local Similarity 23.8%; Pred. No. 0.17;
Matches 64; Conservative 29; Mismatches 79; Indels 97; Gaps 14;
QY 1 MRGSHHHHGGSQIPYTTVNGTSQNILSS-----LTF-NKNQQISYKD----- 43
DB 1 MRGSHHHHGGSGARDISST-NVDTLVSPSKIEDGGKTVKMTFDDKNGKIQNGDMIKVA 59
QY 44 -----IENKVKSV-----LYFN-----RGISDLDLRLSKOAKY 72
DB 60 WPTSTGVKIEGSKVPLTVRGEQVGQAVITPDGATITFNDKVEKLSDV-----SGFAFFE 115
QY 73 VHFKNGTKRVDLKGAGHTADLINTSDIKATSV---NVDTKQVKDKKA----- 118
DB 116 VQGRNLQ-----TNTSDDKVATITSGNKSSTNVVHKSEAGTSVFYKGTG 161
QY 119 -----KANQVQPYTIVNGTSQNILSNLTFKK-----NQQISYKDLNNV-----KSVLKSNR 166
DB 162 DMLPDDTTHVRWFLNNKESYVSKDITIKQIQGGQQLDLSTLNTVGTSHNSYSGQS 221
QY 167 GTTDDVLDLRLSKOAKFTVFNKGTKKVIDL 195
DB 222 AITDFE-KAFPGSKITV---DNTRKTIIV 246

RESULT 21
ABU44119
ID ABU44119 standard; protein; 374 AA.
XX
XX AC ABU44119;
XX
XX DT 19-JUN-2003 (first entry)
XX
XX DE Protein encoded by Prokaryotic essential gene #29646.
XX
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX OS Streptococcus mutans.
XX
XX PN WO200277183-A2.
XX
XX PD 03-OCT-2002.
XX
XX PF 21-MAR-2002; 2002WO-US009107.
XX
XX PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX PA (ELIT-) ELITRA PHARM INC.
XX
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
DR N-PSDB; ACA47989.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 72043; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 374 AA;

Query Match 11.3%; Score 115; DB 6; Length 374;

Best Local Similarity 24.2%; Pred. No. 0.15; Mismatches 63; Indels 58; Gaps 13;

Matches 52; Conservative 42; Mismatches 63; Indels 58; Gaps 13;

QY 18 ITVNGTSQNLSTFNKQOISYKDINKKSVLY-----FNKGISDIDLRLSKQAKYTV 73

DB 131 IIVSGNKNAIESQLI-----BELGIKK-SDYLITLFOANRFRNLKSKD-KWVKEAKLVY 184

QY 74 HFKNQ-TKRVD-----LKAGHTADLINTSDIKA--ISVNVDTKKQVK 114

DB 185 HFPNHFTRVKEVRIIAYRQTDKGYVPILENGTR-VDTVNASELPGSFVTINLDQEKVR 243

QY 115 -----DKEAKANVQVPIITVNGTSQNLSTFNKQOISYKDLENNKSVLXSN 165

DB 244 ELVQKLAKLDKSLVGSIKV--ISSVNSSSTKDLLLEMKDN-----NSRVPL--- 289

QY 166 RGITDVLRLSKQAKFTVNFKNGTKKVIDLKAGIY 200

DB 290 ---SEIDTKLPPYYSIKKILTDGS--IVDMVGIIY 319

RESULT 22

ABB79625

ID ABB79625 standard; protein; 393 AA.

AC ABB79625;

XX

DT 21-OCT-2002 (first entry)

XX

DE E. coli expressed P. falciparum MSP-1 recombinant protein.

XX

KW Merozoite surface protein-1; MSP-1; malaria; vaccine; protozoacide.

XX

OS Plasmodium falciparum.

OS Synthetic.

OS Chimeric.

XX

PN WO200258727-A2.

XX

PD 01-AUG-2002.

XX

PF 25-JAN-2002; 2002WO-US002554.

XX

PR 26-JAN-2001; 2001US-0264535P.

XX

PA (REED-) REED ARMY INST RES WALTER.

XX

PI Lyon JA, Angov E, Cohen JD, Voss G;

XX

WPI; 2002-590798/63.

XX

XX New vaccine comprising Plasmodium falciparum MSP-142 protein and an

PT adjuvant, useful against malaria or for eliciting immune responses

PT against P. falciparum.

XX

PS Disclosure; Page 92-93; 99pp; English.

XX

CC The present sequence is that of a recombinant protein comprising

CC Plasmodium falciparum 3D7 merozoite surface protein-1 42 kDa fragment

CC (MSP-1(42)) with an N-terminal (His)6-tag for affinity purification and 3

CC linker amino acids. The levels of expression of the recombinant protein,

CC encoded by plasmid pT-TEGR-42(AT), were 2-5% of total E. coli protein,

CC and the protein was correctly folded based on immunoreactivity with a

CC series of MSP-1(19) specific monoclonal antibodies. The present invention

CC relates to the production of large amounts of MSP-1(42) which maintain

CC conformational epitopes critical for development of vaccines. The

CC vaccines are useful against malaria or for eliciting immune responses

CC against P. falciparum. The recombinant MSP-1(42) proteins are useful in

CC diagnostic assays, for in vitro monitoring of malaria infection or

CC prognosing the response to treatment of malaria patients, and for

CC production of antibodies used for malaria antigen detection or as

CC therapeutic or prophylactic agents

XX

SQ Sequence 393 AA;

Query Match 11.3%; Score 114.5; DB 5; Length 393;

Best Local Similarity 22.5%; Pred. No. 0.18;

Matches 53; Conservative 34; Mismatches 90; Indels 59; Gaps 10;

QY 4 SHHHHHH-GSQIP--YITVNGTSQNLSTFNKQOISY-----KDIENKV 48

DB 2 AHHHHHHPGSGIEGRGTMAISVTMDNILSG--FENEYDVLYLPLAGVYRSLKKQIE--- 56

QY 49 KSVLYFNRGIDIDLRLSKQAKYTVHFKNGTKRVDLKGHTADLINTSDIKAISVNV 108

DB 57 KNITFNLNLNDILNSRLKRRKYFL-----DVLESDLMQPKHSSNEYIIE 102

QY 109 TKQVKQKQKAKANVQVPIITVNGTSQNI---LSNLTFKKNQOISYKDLENNKSVLXSN 165

DB 103 DSFKLLNSEQNTLLKSYKIKESVENDIKFAQEGISYVEKVLAKYKDDLESIKKVIKEE 162

QY 166 R-----GITDVLRLSKQAKFTVN-----FKNGTKKV----IDLKAGI 199

DB 163 KEKFPSSPPTTPSPAKTDEQKESKFLPLTNTIETLYNLNLNKIDYILNLKAKI 218

RESULT 23

ABP71273

ID ABP71273 standard; protein; 393 AA.

XX

AC ABP71273;

XX

DT 28-APR-2003 (first entry)

XX

DE P. falciparum MSP1_42 (3D7) protein sequence in clone pET42A.

XX

KW MSP-1_42; merozoite protein; protozoacide; vaccine; malaria; mosquito.

XX

OS Plasmodium falciparum.

XX

PN WO2003004525-A2.

XX

PD 16-JAN-2003.

XX

PF 25-JAN-2002; 2002WO-US002428.

XX

PR 29-JAN-2001; 2001US-0264535P.

PR 26-OCT-2001; 2001US-0347584P.

XX

PA (REED-) REED ARMY INST RES WALTER.

XX

PI Lyon JA, Angov E;

XX

WPI; 2003-221577/21.

XX

XX New recombinant Plasmodium falciparum merozoite protein (MSP)-142 which

PT retains its native folding, useful for detecting and preventing malaria

PT infection, and for antibody production.

XX

PS Disclosure; Page 96-97; 104pp; English.

XX

CC The invention relates to a recombinant Plasmodium falciparum merozoite

CC protein, (MSP)-1 42 which retains its native folding. The protein is

CC useful as a diagnostic reagent, in antibody production, and as a vaccine

CC against malaria. The antibody may also be used for detecting and treating

CC chronic malaria infection. The present sequence represents a recombinant

CC E. coli expressed P. falciparum MAP-1_42 protein expressed in clone

CC pET42A

XX

SQ Sequence 393 AA;

Query Match 11.3%; Score 114.5; DB 6; Length 393;

Best Local Similarity 22.5%; Pred. No. 0.18;

PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haseiback R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
DR N-PSDB; ACA37043.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 61097; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 963 AA;

Query Match 10.9%; Score 111; DB 6; Length 963;
Best Local Similarity 21.6%; Pred. No. 1.2;
Matches 51; Conservative 46; Mismatches 73; Indels 66; Gaps 9;
QY 13 QIPYITVN-GTSONILSLTNKKNQIYKDIENKVK----- 49
DB 119 KIETULTVNNKTTLELIAETDLAKQYQLLDNLVNLKLNKANKFOLKEKLSIKAIRAKL 178
QY 50 ----SVLYFNRGISDID-----LRLSKQAKYTVHFKNGTKRVRVDLKAGIHTAD----- 93
DB 179 NKELNILYEN-AVSSQDKKQPELTISEQIYQYENLYINNQRVIAVHQKLNALQLKKIYK 237
QY 94 ----LIINTSDIKAISVNVDTKKQVXDKEAKANVQVPYITVNGTSQNLISNLTFFKN--- 146
DB 238 TNILLNLLNLDTKALQASIDNRYEVINOYS-----AIESSIQTILNSLTNQANLNV 287
QY 147 ---QQISKYDLENVNVKSVLKNRGITDVLRLSKQAKYTVHFKNGTKRVRVDLKAGI 199
DB 288 EFGQLQSVSLQNDVKTLEKET---SNQKILSKNLE---DYQNQLKKQISTRQSL 337

RESULT 28
AAW26306
ID AAW26306 standard; peptide; 31 AA.
XX
XX AC AAW26306;
XX
XX 17-OCT-2003 (revised)
DT 16-NOV-1997 (first entry)
XX
XX Staphylococcal MHC class II antigen analogue repeat subsegment 5.
XX
XX MHC II analogue protein; major histocompatibility complex;
KW staphylococcal; adhesin; virulence factor; vaccine.
XX
XX Staphylococcus aureus; strain FDA 574.
XX
XX OS US5648240-A.
XX
XX PD 15-JUL-1997.
XX
XX PF 24-MAY-1994; 94US-00248021.
XX
XX PR 24-MAY-1994; 94US-00248021.
XX
XX (TEXA) UNIV TEXAS A & M.
XX
XX Jonsson K, Hook M, Patti JM, Gurusiddappa S;
XX WPI; 1997-372059/34.
XX
XX DNA encoding Staphylococcus aureus broad spectrum adhesin - for
PT production of recombinant adhesin for use in vaccines.
XX
XX Claim 17; Col 15-16; 30pp; English.
XX
XX Claimed peptides (AAW26302-07) correspond to subsegments of repeat units
CC of a broad-spectrum adhesin (see AAW26301) of Staphylococcus aureus FDA.
CC These subsegments show high homology with eukaryotic MHC Class II
CC antigens. They can be used in claimed compositions to induce an immune
CC response to *S. aureus* in an animal. (Updated on 17-OCT-2003 to
CC standardise OS field)
XX
XX SQ Sequence 31 AA;
Query Match 10.8%; Score 110; DB 2; Length 31;
Best Local Similarity 66.7%; Pred. No. 0.014; Mismatches 6; Indels 0; Gaps 0;
Matches 20; Conservative 20; Mismatches 6; Indels 0; Gaps 0;
QY 60 DIDLRLSKQAKYTVHFKNGTKRVRVDLKAGI 89
DB 1 DIELKFAKQAKYTIHFKNGTQVVDLKSDI 30

RESULT 29
ADG32048
ID ADG32048 standard; protein; 438 AA.
XX
XX AC ADG32048;
XX
XX 26-FEB-2004 (first entry)
XX
XX Mutant B_licheniformis secreted polypeptide SeqID 18.
XX
XX mutant; host cell; production yield; shelf life; product stability;
KW purity; secreted; mutein.
XX
XX Synthetic.
OS Bacillus licheniformis.
XX
XX WO2003093453-A2.
XX

Db 545 IGGNISQKEGNTLTISSDKININQITIKKGVNKEDSDSSSTANNANLTIKTKELQLTGDLN 604

QY 175 LSKQAKFTVNFNGTKKVI 193

Db 605 ISGFDKAEITAKEGADLII 623

RESULT 31

AAB01834

ID AAB01834 standard; protein; 1101 AA.

AC AAB01834;

XX

DT 12-SEP-2003 (revised)

DT 11-SEP-2000 (first entry)

XX

DE Haemophilus influenzae strain LDC23 HMW1A protein, SEQ ID NO:43.

XX

KW HMW protein; hmw gene; hmwA1; hmwA2; high molecular weight;

KW non-typeable Haemophilus influenzae; NTHi; non-encapsulated;

KW recombinant production; Escherichia coli; antibacterial; vaccine;

KW human disease; otitis media; epiglottitis; pneumonia; tracheobronchitis;

KW detection; diagnosis.

XX

OS Haemophilus influenzae; strain LDC23.

XX

FN WO200020609-A2.

XX

PD 13-APR-2000.

XX

PF 07-OCT-1999; 99WO-CA000938.

XX

PR 07-OCT-1998; 98US-00167568.

PR 08-DEC-1998; 98US-00206942.

XX

PA (CONN-) CONNAUGHT LAB LTD.

XX

PI Loosmore SM, Yang Y, Klein MH;

XX

DR WPI; 2000-303789/26.

DR N-PSDB; AAA52183.

XX

PT Nucleic acid molecule for producing recombinant high molecular weight

PT proteins of Haemophilus which are used as a vaccine to provide protection

PT against Haemophilus induced diseases in humans.

XX

PS Claim 12; Fig 22A-P; 307pp; English.

XX

CC The invention relates to the recombinant production of Haemophilus

CC influenzae high molecular weight (HMW) proteins in Escherichia coli. The

CC expression construct used to effect recombinant expression comprises a

CC promoter functional in E. coli (e.g., the T7 promoter) operably linked to

CC a modified hmwABC operon from a non-typeable (non-encapsulated) H.

CC influenzae (NTHi). Most HMW-expressing NTHi strains contain two hmw gene

CC clusters termed hmwIABC and hmw2ABC. Each hmwABC operon comprises hmwA,

CC hmwB and hmwC genes. The hmwA genes encode the structural HMWA proteins

CC and the hmwB and hmwC genes encode accessory proteins which are

CC responsible for post-translational processing and secretion of the HMWA

CC proteins. The modified hmwABC operon used in the expression construct of

CC the invention contains an A gene modified such that it encodes only the

CC mature HMWA. The invention also discloses hmwA genes (AAA52175-A52198)

CC and HMWA proteins (AAB01824-B01849) from the non-typeable H. influenzae

CC strains Joysc, Ki, K21, LDC22, PMH1, 15 and 12. The nucleic acids and

CC vectors are used for the production of recombinant H. influenzae HMW

CC proteins which can be used as vaccines to mediate a humoral or cell-

CC mediated immune response to provide protection against diseases in humans

CC caused by H. influenzae (e.g., otitis media, epiglottitis, pneumonia and

CC tracheobronchitis). The HMW proteins are also useful as antigens in

CC immunoassays for detecting antibodies against Haemophilus, HMW proteins

CC and/or HMW peptides. The nucleotide sequences encoding the HMW proteins

CC can be used to isolate and clone hmw genes from other non-typeable

CC strains of Haemophilus via hybridisation reactions. The present sequence

CC represents an HMWA protein from a non-typeable strain of H. influenzae.

CC (Updated on 12-SEP-2003 to standardise OS field)

XX SQ Sequence 1101 AA;

Query Match 10.8%; Score 110; DB 3; Length 1101;
 Best Local Similarity 25.1%; Pred. No. 1.8;
 Matches 50; Conservative 29; Mismatches 76; Indels 44; Gaps 8;

QY 18 ITVNGTSQNILSSLT-----FNKNQOISYKDIENKVKSVLYFNRGISIDILRLSKQAKY 71

Db 452 LTVTGSAINIEKULTVEGSAKFLANPNYSFN-----VSGLPDNGQKSNISI-----AKG 500

QY 72 TVHFK--NGTKRVVDLKAGIHTADLINTSD-----IKAISVNVDTKKQVKDKAKANVQ 123

Db 501 GAHFKDINNTK-----SLNITNTSDSAYRTIIEGNTITNSGDLNITDNKNAETQ 550

QY 124 VPYIT-----TVNGTSQNILSLNLTFFK--NQOISYKDLNENKSVLSKSNRGITDVLDR 174

Db 551 IGGNISQKEGNTLTISSDKININQITIKKGVNKEDSDSSSTANNANLTIKTKELQLTGDLN 610

QY 175 LSKQAKFTVNFNGTKKVI 193

Db 611 ISGFDKAEITAKEGADLII 629

RESULT 32

AAB01825

ID AAB01825 standard; protein; 1221 AA.

AC AAB01825;

DT 12-SEP-2003 (revised)

DT 11-SEP-2000 (first entry)

XX Haemophilus influenzae strain Joysc mature HMW1A protein, SEQ ID NO:28.

XX Mature HMW protein; hmw gene; hmwA1; hmwA2; high molecular weight;

XX non-typeable Haemophilus influenzae; NTHi; non-encapsulated;

XX recombinant production; Escherichia coli; antibacterial; vaccine;

XX human disease; otitis media; epiglottitis; pneumonia; tracheobronchitis;

XX detection; diagnosis.

XX Haemophilus influenzae; strain Joysc.

XX WO200020609-A2.

PD 13-APR-2000.

PF 07-OCT-1999; 99WO-CA000938.

PR 07-OCT-1998; 98US-00167568.

PR 08-DEC-1998; 98US-00206942.

PA (CONN-) CONNAUGHT LAB LTD.

PI Loosmore SM, Yang Y, Klein MH;

XX WPI; 2000-303789/26.

XX N-PSDB; AAA52176.

PT Nucleic acid molecule for producing recombinant high molecular weight

PT proteins of Haemophilus which are used as a vaccine to provide protection

PT against Haemophilus induced diseases in humans.

XX Claim 8; Fig 18A-R; 307pp; English.

XX The invention relates to the recombinant production of Haemophilus

XX influenzae high molecular weight (HMW) proteins in Escherichia coli. The

XX expression construct used to effect recombinant expression comprises a

XX promoter functional in E. coli (e.g., the T7 promoter) operably linked to

XX a modified hmwABC operon from a non-typeable (non-encapsulated) H.

XX influenzae (NTHi). Most HMW-expressing NTHi strains contain two hmw gene

XX clusters termed hmwIABC and hmw2ABC. Each hmwABC operon comprises hmwA,

CC hmWB and hmWC genes. The hmWA genes encode the structural HMWA proteins
CC and the hmWB and hmWC genes encode accessory proteins which are
CC responsible for post-translational processing and secretion of the HMWA
CC proteins. The modified hmWABC operon used in the expression construct of
CC the invention contains an A gene modified such that it encodes only the
CC mature HMWA. The invention also discloses hmWA genes (AA52175-A52198)
CC and HMWA proteins (AAB01824-B01849) from the non-typeable H. influenzae
CC strains JcyC, K1, K21, LCDC2, PMH1, 15 and 12. The nucleic acids and
CC vectors are used for the production of recombinant H. influenzae HMW
CC proteins which can be used as vaccines to mediate a humoral or cell-
CC mediated immune response to provide protection against diseases in humans
CC caused by H. influenzae (e.g., otitis media, epiglottitis, pneumonia and
CC tracheobronchitis). The HMW proteins are also useful as antigens in
CC immunoassays for detecting antibodies against Haemophilus, HMW proteins
CC and/or HMW peptides. The nucleotide sequences encoding the HMW proteins
CC can be used to isolate and clone hmw genes from other non-typeable
CC strains of Haemophilus via hybridisation reactions. The present sequence
CC represents a mature HMWA protein from a non-typeable strain of H.
CC influenzae. (Updated on 12-SEP-2003 to standardise OS field)
XX
SQ Sequence 1221 AA;
Query Match 10.8%; Score 110; DB 3; Length 1221;
Best Local Similarity 23.7%; Pred. No. 2.1;
Matches 54; Conservative 28; Mismatches 82; Indels 64; Gaps 9;
QY 3 GSHHHHGGSQIPYTI-----TVNGTSQNILSSLTFFKNQOISYKDINKVKS 51
DB 422 GSDFDNH---QKPLTIKKDVIINSNLTAGGNVINGNLTGVNNGANL--KAITNFTFN 476
QY 52 --LYFNRGISDIDLRLSKQAKYTVHFKNGTGRVVDLKGAIHTADLINTSDIKAISVNDT 109
DB 477 GGLFDNKGNSNISI-----ARGGAKPKDINNNTSSLN--ITTNSDT 514
QY 110 KKQ-----VKDKEAKANVQVPYTI-----TVNGTSQNILSSLTFFK-- 145
DB 515 TYRTIIEGNTWKAGDLNIIDNKGNAEIQIGNISQKEGNLTISDDKININQITIKGV 574
QY 146 NQOISYKDLNENNVKSVLKSNGRIGTVDLRLSKQAKFTVNFKNGTKKVI 193
DB 575 NKEDSDSSSTANNANLTIKTKELQTLGDLNISGFDKAEITAKEGADLII 622
RESULT 33
ID AAB01824
XX AAB01824 standard; protein; 1227 AA.
AC AAB01824;
XX 12-SEP-2003 (revised)
DT 11-SEP-2000 (first entry)
XX Haemophilus influenzae strain JcyC HMWA protein, SEQ ID NO:26.
XX HMW protein; hmw gene; hmwa2; high molecular weight;
KW non-typeable Haemophilus influenzae; NTHi; non-encapsulated;
KW recombinant production; Escherichia coli; antibacterial; vaccine;
KW human disease; otitis media; epiglottitis; pneumonia; tracheobronchitis;
KW detection; diagnosis.
XX Haemophilus influenzae; strain JcyC.
OS
XX
XX WO2000020609-A2.
PN
XX
PD 13-APR-2000.
XX
PF 07-OCT-1999; 99WO-CA000938.
XX
XX 07-OCT-1998; 98US-00167568.
PR 08-DEC-1998; 98US-00206942.
XX
XX (CONN-) CONNAUGHT LAB LTD.
PA
XX

PI Loosmore SM, Yang Y, Klein MH;
XX WPI; 2000-303789/26.
DR N-PSDB; AA52175.
XX
PT Nucleic acid molecule for producing recombinant high molecular weight
PT proteins of Haemophilus which are used as a vaccine to provide protection
PT against Haemophilus induced diseases in humans.
XX
PS Claim 12; Fig 18A-R; 307pp; English.
XX
CC The invention relates to the recombinant production of Haemophilus
CC influenzae high molecular weight (HMW) proteins in Escherichia coli. The
CC expression construct used to effect recombinant expression comprises a
CC promoter functional in E. coli (e.g., the T7 promoter) operably linked to
CC a modified hmWABC operon from a non-typeable (non-encapsulated) H.
CC influenzae (NTHi). Most HMW-expressing NTHi strains contain two hmw gene
CC clusters termed hmWABC and hmW2ABC. Each hmWABC operon comprises hmWA,
CC hmWB and hmWC genes. The hmWA genes encode the structural HMWA proteins
CC and the hmWB and hmWC genes encode accessory proteins which are
CC responsible for post-translational processing and secretion of the HMWA
CC proteins. The modified hmWABC operon used in the expression construct of
CC the invention contains an A gene modified such that it encodes only the
CC mature HMWA. The invention also discloses hmWA genes (AA52175-A52198)
CC and HMWA proteins (AAB01824-B01849) from the non-typeable H. influenzae
CC strains JcyC, K1, K21, LCDC2, PMH1, 15 and 12. The nucleic acids and
CC vectors are used for the production of recombinant H. influenzae HMW
CC proteins which can be used as vaccines to mediate a humoral or cell-
CC mediated immune response to provide protection against diseases in humans
CC caused by H. influenzae (e.g., otitis media, epiglottitis, pneumonia and
CC tracheobronchitis). The HMW proteins are also useful as antigens in
CC immunoassays for detecting antibodies against Haemophilus, HMW proteins
CC and/or HMW peptides. The nucleotide sequences encoding the HMW proteins
CC can be used to isolate and clone hmw genes from other non-typeable
CC strains of Haemophilus via hybridisation reactions. The present sequence
CC represents an HMWA protein from a non-typeable strain of H. influenzae.
CC (Updated on 12-SEP-2003 to standardise OS field)
XX
SQ Sequence 1227 AA;
Query Match 10.8%; Score 110; DB 3; Length 1227;
Best Local Similarity 23.7%; Pred. No. 2.1;
Matches 54; Conservative 28; Mismatches 82; Indels 64; Gaps 9;
QY 3 GSHHHHGGSQIPYTI-----TVNGTSQNILSSLTFFKNQOISYKDINKVKS 51
DB 428 GSDFDNH---QKPLTIKKDVIINSNLTAGGNVINGNLTGVNNGANL--KAITNFTFN 482
QY 52 --LYFNRGISDIDLRLSKQAKYTVHFKNGTGRVVDLKGAIHTADLINTSDIKAISVNDT 109
DB 483 GGLFDNKGNSNISI-----ARGGAKPKDINNNTSSLN--ITTNSDT 520
QY 110 KKQ-----VKDKEAKANVQVPYTI-----TVNGTSQNILSSLTFFK-- 145
DB 521 TYRTIIEGNTWKAGDLNIIDNKGNAEIQIGNISQKEGNLTISDDKININQITIKGV 580
QY 146 NQOISYKDLNENNVKSVLKSNGRIGTVDLRLSKQAKFTVNFKNGTKKVI 193
DB 581 NKEDSDSSSTANNANLTIKTKELQTLGDLNISGFDKAEITAKEGADLII 628
RESULT 34
ID ABP39160
XX ABP39160 standard; protein; 288 AA.
AC ABP39160;
XX
DT 24-JUL-2002 (first entry)
XX
DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:4005.
XX
KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW antibacterial; gene therapy.

[illegible]

```
CC prodn. of antimalarial vaccines. See also AAQ13727-8
XX Sequence 493 AA;
SQ
  Query Match      10.6%; Score 107.5; DB 2; Length 493;
  Best Local Similarity 25.4%; Pred. No. 1;
  Matches 50; Conservative 24; Mismatches 72; Indels 51; Gaps 9;
QY 1 MRGSHHHHGGSQIPYITVNGTSQ--NILSSLTFNKNQOISYKDIENKVKSVLYFNRI 58
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 MRGSHHHHGGSS-----VNSVKENNNKVNNAVHKI NAVDKNAVNKNSVKNLN--VV 52
QY 59 SDIDLRLSKQAKY---TVHFKNGTGRVVDLKGAIHTADLINTSDIKASVNVDTKKQVKD 115
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 53 NKTNVLSKLNNAVYKNSVHKMNAVNKV-----NAVKNVNAVN-----KVNVNKKDILN 101
QY 116 KEAKANVQVPYITVNGTSQNILSLTFNKNQOISYKDIENKVKSVLSNRGITDVLRL 175
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 102 K-----LNALYKNNAVYKNNAL--NKVSAVNK-----VSAVNKVS 134
QY 176 SKQAKFTVNFKNGTKKV 192
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 135 AVNKMGA VNRVNGVNVK 151
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 38
AAW31555
ID AAW31555 standard; protein; 345 AA.
XX
AC AAW31555;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 21-MAY-1998 (first entry)
XX
DE Fibronectin-binding MSCRAMM derivative pCF33.
XX
KW Fibronectin; pCF33; collagen binding protein; sepsis; infection;
KW microbial surface component recognising adhesive matrix molecule; MSCRAMM;
KW adhesin; vaccine; immunisation; diagnosis; therapy; mastitis.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT Peptide 1..12
  /note= "vector pQE30-derived peptide"
XX
XX WO9743314-A2.
XX
XX 20-NOV-1997.
XX
XX 14-MAY-1997; 97WO-US008210.
XX
XX 16-MAY-1996; 96US-0017678P.
XX
XX (TEXA ) UNIV TEXAS A & M SYSTEM.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Hoeek M, Patti JM, House-Pompeo K, Sthanam N, Symersky J;
XX WPI, 1998-008801/01.
XX
XX Antibody that interacts with collagen binding domain of Staphylococcal
XX cna gene product - useful to prevent bacterial sepsis in animal infected
XX with Staphylococcus aureus.
XX
XX Disclosure; Page 91; 143pp; English.
XX
XX This protein comprises Staphylococcus aureus fibronectin-binding
XX microbial surface component recognising adhesive matrix molecule (MSCRAMM)
XX derivative pCF33, plus a vector-derived N-terminal peptide. The invention
XX relates to claimed nucleic acid sequences (see AAT93436-38) encoding S.
XX aureus collagen binding protein (CBP) epitopes M17, M31 and M55 (see
```

```
CC AAQ31552-54) that confer protection against S. aureus infection. CBP
CC protein and antigenic epitopes are contemplated for use in the treatment
CC of pathological infections, especially to prevent bacterial adhesion to
CC collagen. The epitopes are also contemplated for use in the preparation
CC of vaccines and as carrier proteins in vaccine formulations, as well as
CC in the formulation of compositions for the prevention of S. aureus
CC infection. pCF33 and pQD (see AAW31556) were used to raise anti-MSCRAMM
CC polyclonal antibodies used in passive immunisation against bovine
CC mastitis. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 27-
XX AUG-2003 to correct OS field.)
XX
SQ Sequence 345 AA;
  Query Match      10.4%; Score 106; DB 2; Length 345;
  Best Local Similarity 23.3%; Pred. No. 0.83;
  Matches 52; Conservative 35; Mismatches 90; Indels 46; Gaps 11;
QY 1 MRGSHHHHGGSQIPYITVNGTS--QNTLSSLTFNKNQOIS--YKDIENKVKSVLYP--N 55
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 MRGSHHHHGGSSVMAADAPAAAGTDITNQTNTVTGIDSGTTVYPHQAQGVKLNLYGFSV 60
QY 56 RGISDIDLRLSKQAKYTVHFKNGTGRVVDLKGAG-----IHT-ADLINT-S 98
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 61 SAVKGDTEKIIVPKELNLNGVTSTAKVPIAGDQVLANGVIDSGNVITFTDVTYTKD 120
QY 99 DIKAI-----SVNVDTKKQVKQKKEAKANYQVP-----YTITVNGTSQNI 137
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 121 DVKATLTMPAVIDPENVKKTGNVTLATGIGSTTANKTVLVVDYKYGKFNLSIKGTIDQI 180
QY 138 -LSNLTFFKNQOISYKDIENKVKSVLSKN-RGITDVLRLSKQ 178
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 181 DKTNTYRTQTYVNPSPG-DNVIAPVLTCNLKPTDSDNALIDQQ 222
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 39
ABM79015
ID ABM79015 standard; protein; 560 AA.
XX
AC ABM79015;
XX
DT 15-JAN-2004 (first entry)
XX
DE Staphylococcus epidermidis SdrG NIN2N3 domain.
XX
KW SdrG; surface protein; infection; antibacterial; vaccine.
XX
OS Staphylococcus epidermidis.
XX
XX WO2003076470-A1.
XX
XX 18-SEP-2003.
XX
XX 05-MAR-2003; 2003WO-US006415.
XX
XX 05-MAR-2002; 2002US-0361324P.
XX
XX (INH-) INHIBITEX INC.
XX
XX (TEXA ) UNIV TEXAS A & M SYSTEM.
XX
XX Patti JM, Hutchins JT, Hall A, Domanski P, Patel P, Hook M;
XX Robbins J, Vernachio J, Bowden MG;
XX WPI; 2003-722324/68.
XX
XX N-PSDB; ACF0624.
XX
XX New antibody recognizing a Staphylococcus epidermidis protein comprising
XX SdrG NIN2N3, SdrG N2N3 or SdrGR2 useful for preparing a composition for
XX treating or preventing a coagulase-negative Staphylococcal infection.
XX
XX Claim 27; Page 24-25; 78pp; English.
XX
XX The present sequence comprises the protein sequence of the NIN2N3 region
XX (amino acids 50-597), or putative A domain, of the SdrG surface protein
CC
```

